

The repair of DNA breaks in *Escherichia coli*  
and analysis of the bacterial recombination  
protein RecN

Stuart Robert Wood, B.Sc. (Hons.)

A thesis submitted to the University of Nottingham for the  
degree of Doctor of Philosophy,  
September 2008.



## Table of Contents

Abbreviations .....	iv
Acknowledgments .....	vi
Statement .....	vii
Abstract .....	viii
<b>Chapter 1. Introduction .....</b>	<b>1</b>
1.1 DNA damage and repair .....	1
1.2 Oxidative damage, damage prevention and repair .....	7
1.3 DSB formation.....	9
1.4 DSB repair.....	14
1.5 RecN.....	25
1.6 Summary .....	38
<b>Chapter 2. Materials and methods .....</b>	<b>40</b>
2.1 Materials .....	40
2.2 Biological methods .....	55
2.3 DNA preparation and analysis.....	62
2.4 Protein preparation and analysis.....	66
<b>Chapter 3. The essential nature of the OxyR regulon in the absence of recombination.....</b>	<b>72</b>
3.1 Identification of genes that are synthetically lethal with the recombination genes <i>recB</i> and <i>recN</i> .....	73
3.2 Mutations in the OxyR regulon are synthetically lethal with <i>rec</i> mutations.....	78
3.3 Discussion .....	85
<b>Chapter 4. RecN homologs can function in <i>E. coli</i>.....</b>	<b>89</b>
4.1 RecN homology .....	89
4.2 Heterologous genetics .....	94
4.3. <i>H. influenzae</i> RecN can function instead of <i>E. coli</i> RecN <i>in vivo</i> .....	100
4.4 Discussion.....	105

<b>Chapter 5 Purification of RecN .....</b>	<b>108</b>
5.1 Expression of RecN proteins in <i>E. coli</i> .....	109
5.2 Purification of RecN from <i>H. influenzae</i> .....	111
5.3 Purification of the other RecN homologs .....	120
5.4 Discussion .....	123
<b>Chapter 6 The <i>in vitro</i> activities of RecN .....</b>	<b>125</b>
6.1 RecN has a detectable ATPase activity .....	126
6.2 RecN requires ATPase activity to function.....	137
6.3 DNA Binding Activities of RecN.....	139
6.4 Structural properties of RecN.....	145
6.5 Discussion .....	154
<b>Chapter 7 General discussion.....</b>	<b>160</b>
7.1 Suggested roles of RecN.....	162
7.2 RecN and SMC proteins can aid accurate genome transmission.....	166
7.3 RecN and SMC proteins are involved in DNA replication .....	170
7.4 Models of RecN function .....	171
7.5 Future work.....	179
<b>Chapter 8 References .....</b>	<b>182</b>

# Abbreviations

Aq	prefix for <i>Aquifex aeolicus</i> proteins
ADP	adenosine diphosphate
amp	ampicillin
ATP	adenosine triphosphate
apra	apramycin
Bf	prefix for <i>Bacteriodes fragilis</i> proteins
bp	base pair
Bs	prefix for proteins from <i>Bacillus subtilis</i>
BER	base excision repair
BIR	break induced replication
BSA	bovine serum albumin
cat	chloramphenicol acetyl transferase
ChIP	chromosome immune-precipitation
CTAB	cetyl-triethylammonium bromide
DNA	deoxyribonucleic acid
DSB	double-strand break
DAPI	4',6' – Diamidino-2-phenylindole
dH <sub>2</sub> O	distilled water
dsDNA	double-stranded deoxyribonucleic acid
DTT	dithiothreitol
Ec	prefix for <i>Escherichia coli</i> proteins
EDTA	diaminoethanetetra-acetic acid
GFP	green fluorescent protein
Hi	prefix for <i>Haemophilus influenzae</i> proteins
HJ	Holliday Junction
HR	homologous Recombination
IPTG	isopropylthiogalactoside
IR	ionising radiation
kan	kanamycin
MMC	mitomycin C



MMR	mismatch repair
nt	nucleotides
NER	nucleotide excision repair
NTP	nucleotide triphosphate
NHEJ	non-homologous end joining
OD	optical density
Pi	inorganic phosphate
PCR	polymerase chain reaction
pfu	plaque forming units
PAGE	polyacrylamide gel electrophoresis
<i>rec</i> <sup>+</sup>	recombination proficient
<i>rec</i> <sup>-</sup>	recombination deficient
RNA	ribonucleic acid
RNI	reactive nitrogen intermediates
ROS	reactive oxygen species
SDS	sodium dodecyl sulphate
SEM	scanning electron microscope
SMC	structural maintenance of chromosomes
SOD	superoxide dismutase
SSB	single strand binding protein
SDSA	synthesis-dependent strand annealing
ssDNA	single-stranded deoxyribonucleic acid
tc	tetracycline
tm	trimethoprim
TEM	transmission electron microscope
UV	ultra violet
w/v	weight per volume
v/v	volume per volume
X-gal	5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside
YFP	yellow fluorescent protein

# Acknowledgments

Firstly I would like to take this opportunity to thank my supervisor Professor Robert Lloyd for his advice, encouragement and enthusiasm for the project throughout my PhD. I am deeply indebted to all those in the laboratory, past and present, for their advice and assistance. A special thank you to Geoff and Jane, not only for sharing your knowledge and experience throughout my PhD, but also your kind assistance and tireless patience in teaching me proper grammar and english in general. I would also like to acknowledge the work done in conjunction with Jane Grove during this project, much of the heterologous strain construction was performed by Dr. Jane Grove (University of Nottingham) based on expression clones I initially produced especially those strains that were used in the I-SceI viability assays, those strains produced by Jane are easily denoted by the JIG prefix. Work on this project has since continued under Jane's guidance

Thank you to Akeel, Tim, Andy, Ed, Christian, Thorsten and Stéphane for sharing your knowledge and experience. To my fellow PhD students, Amy, Jing, Jing, Belinda, Colin, Sam and Michelle, for encouragement and support, it's nice knowing you're not alone. To all those in the lab, thank you for making my time here an enjoyable one. I would also like to thank Carol and Lynda for their superb technical support and I am deeply appreciative of the assistance provided by those outside of the laboratory, whether it be providing materials Dr. G. Thomas (University of York) or technical assistance and advice Karen, Katy and Jodi thank you all.

Thank you also goes to all those outside the lab and the friends I have made in the department and around Nottingham. For the support, amusement and entertainment you have provided, you have made my time here memorable. Thank you to my family, especially my parents and grandparents your support has not gone unnoticed and is deeply appreciated. Finally, thank you to Carly, your love and support has been invaluable, as well as the proof reading, thank you.

## **Statement**

The work presented here was carried out solely by the author (unless cited) since October 2004 in the Institute of Genetics, University of Nottingham, whilst registered as a full-time Ph.D. student. This thesis is comprised of original work, which has not been presented for examination in any other form. This research was supported by the MRC and BBSRC. Whilst carrying out the work described within this thesis, I was fully aware of the hazards associated with the materials and techniques I was using, as advised in the Control of Substances Hazardous to Health regulations. The guidelines laid down in these regulations, and departmental rules, were adhered to at all times. I was also aware of, and followed, the regulations concerning the use and disposal of radioisotopes.

# Abstract

A DNA double-strand break is an exceptionally toxic lesion that threatens the structural and functional integrity of the genome. In this thesis the repair of DSBs was investigated using the bacterium *Escherichia coli*, which repairs DNA breaks almost exclusively by homologous recombination. The studies described focus on the repair of damage induced by reactive oxygen species, but especially on the RecN protein, which is associated specifically with the repair of double-strand breaks.

The RecN protein is highly conserved across bacterial species and in *E. coli* has been identified as a key factor in the repair of DNA breaks. In this thesis three RecN homologs were analysed. RecN from *Haemophilus influenzae* is shown to be capable of replacing the functions of *E. coli* RecN *in vivo*. However, homologs from *Aquifex aeolicus* and *Bacteriodes fragilis* cannot do so.

Biochemical analysis of all three RecN homologs was undertaken. The *H. influenzae* RecN and *A. aeolicus* RecN were shown to have weak ATPase activity and an ability to interact with single-stranded DNA. ATPase deficient mutants of the RecN proteins were created and used to demonstrate the functional importance of the ATP hydrolysis. In the case of *E. coli* and *H. influenzae*, the ATPase defective mutants failed to function *in vivo*. *In vitro*, the ATPase deficient *H. influenzae* RecN mutant and a similar mutant of *A. aeolicus* RecN failed to interact with single-stranded DNA. These data are discussed in terms of a relationship between RecN and the structural maintenance of chromosome family of proteins.

Finally, a model for RecN activity is presented based on those developed to explain the function of structural maintenance of chromosome proteins and the new data presented here. In this model, RecN is suggested to trap DNA molecules holding a break site and repair template in close proximity, facilitating the repair of DNA breaks by homologous recombination. The possibility of RecN acting as a global, damage induced cohesin is also discussed.

# Chapter 1

## Introduction

The accurate replication of DNA and faithful transmission of duplicated chromosomes are major challenges that dividing cells must meet in order to maintain genomic stability. To address these issues, organisms have evolved numerous mechanisms to preserve the basic integrity of the genome (Giraud *et al.* 2001). Damage to DNA is a major and persistent threat. It can range from simple replication errors that create DNA base pair mismatches, through lesions that affect the deoxyribose sugar backbone, or the nitrogenous bases, all of which can be repaired using the complementary strand as a template, to the extremely genotoxic double-strand break (DSB), where both sugar phosphate chains of the DNA have been broken. This introduction provides an overview of DNA damage and its repair, focusing in particular on oxidative damage to DNA and the repair of DNA breaks.

### 1.1 DNA damage and repair

The DNA in all living cells is constantly under assault and the sources of DNA damage can be split crudely into those that are from external sources (exogenous DNA damage) and those that come from within the cell (endogenous DNA damage). Major sources of exogenous DNA damage include UV light, which induces the formation of pyrimidine dimers and other lesions and with high doses is capable of causing DSBs (Bonura and Smith 1975) (Table 1). DSBs can also be generated by exposure to ionising radiation (IR), which includes both gamma and X-rays (Ward 1975). Chemical mutagens and antibiotics are also known to cause damage to DNA. Many are exploited experimentally, including bleomycin, methyl methanesulphonate and mitomycin C (Iyer & Szybalski, 1963; Povirk *et al.*, 1977).

Endogenous sources cause the majority of DNA damage. In a mammalian cell they can cause 10,000 or more lesions a day (Lindahl 1993; Mitchell *et al.* 2003; Setlow 2001). The principal endogenous cause of damage

are reactive oxygen species (ROS), namely the superoxide anion ( $O_2^{\bullet-}$ ), the hydroxyl radical ( $HO^{\bullet}$ ) and hydrogen peroxide ( $H_2O_2$ ). They arise primarily due to electron leakage from the respiratory chain during oxidative metabolism, but also as a result of water radiolysis by UV and IR (Beckman and Ames 1998; Storz and Imlay 1999; Seaver and Imlay 2004). ROS cause a wide variety of damage to cellular components, including numerous lesions to DNA and their prevalence means that DNA damage is inevitable. Fortunately, cells have evolved numerous repair pathways to correct lesions in general and thus limit the accumulation of potentially mutagenic or genotoxic damage (Table 1.1).

Several pathways exist that can remove those lesions that affect just one strand of the DNA. These pathways act by excision of the strand region affected by the lesion and then make use of DNA polymerases and DNA ligase to synthesise the missing section and restore the DNA to its duplex form (Figure 1.1). Base excision repair (BER) involves the specific recognition of a lesion by a DNA glycosylase, which breaks the N-glycosidic bond connecting the base to the sugar phosphate backbone, followed by cleavage of the backbone by an AP endonuclease (Krokan *et al.* 1997; Lindahl and Wood 1999). Nucleotide excision repair (NER) and the mismatch repair (MMR) pathways both excise much larger sections. NER is a more versatile pathway than BER, as it targets a lesion by recognising the distortions it causes in the DNA rather than the lesion itself and NER can be targeted to actively transcribed regions of the genome, reducing the risk of mutations arising within expressed genes (Friedberg 1985; Sancar 1996a; Wood 1996). Mismatch repair (MMR) allows a cell to remove bases that have been misincorporated during DNA replication. This requires the cell to distinguish between the parental (template) strand containing the correct nucleotide and the newly synthesised strand, containing the misincorporated nucleotide. In the bacterium *Escherichia coli* this is achieved by comparing the methylation state of the two strands (Modrich 1991; Fishel and Kolodner 1995; Kolodner 1996).

Some lesions can be repaired directly, i.e. by reversal of the disfiguring event. Thus, a single-strand nick can be ligated by DNA ligase (Barnes *et al.*

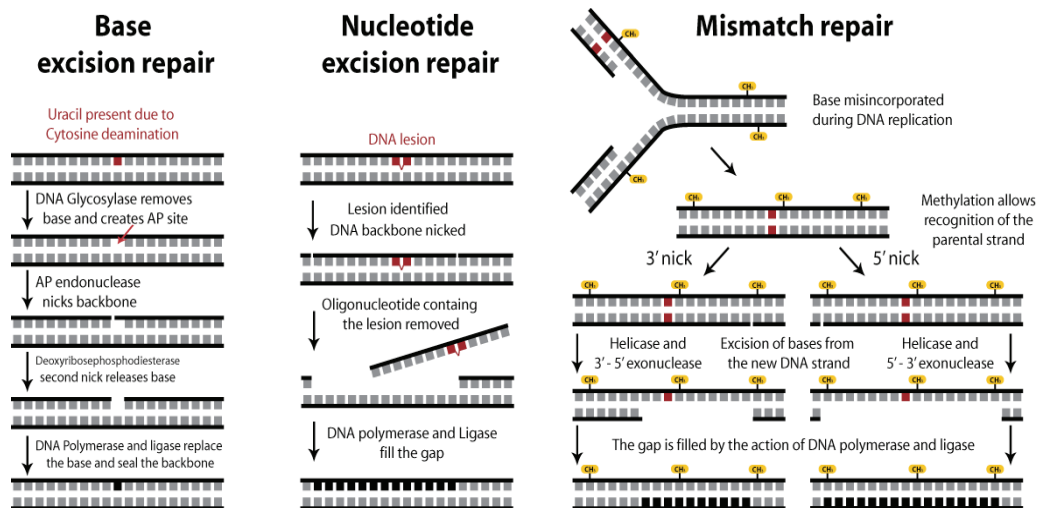
1992), while UV-induced intra-strand cross-links, such as cyclobutane pyrimidine dimers and 6-4 photoproducts, can be broken by the action of a photolyase (Kelner 1949; Sancar 1996b; Sancar 2000). Even chemical modifications of the DNA, such as methylation caused by alkylating agents, can be repaired by protein-mediated removal of the methyl group. In many bacteria this is part of the 'adaptive response' in which exposure to alkylating agents induces expression of the Ada transcription factor. Once activated by methylation Ada itself can remove methyl groups and also induce expression of the AlkB enzyme, which can remove methyl groups from modified bases by oxidative decarboxylation (Samson and Cairns 1977; Lindahl *et al.* 1988; Trewick *et al.* 2002). All of these systems act to restore the DNA to its correct, stable chemical form.

Many lesions prevent synthesis by the replicative DNA polymerase, but have little or no effect on the progression of the replicative helicase and associated replisome components. This can result in gaps being left opposite a lesion, making excision repair impossible. However, most cells possess alternative DNA polymerases that can be recruited to the replisome to catalyse translesion synthesis. These polymerases tend to have low processivity and are displaced by the replicative polymerase as soon as a lesion is bypassed. Hence, they allow DNA synthesis to continue, preserving the DNA duplex so that excision repair can remove the lesion after the replisome has passed. Consequently translesion synthesis allows a lesion to be tolerated, but can lead to misincorporation, resulting in mutation of the genome (Lehmann 2006; Shcherbakova and Fijalkowska 2006). Because of the risk of mutation translesion synthesis is usually thought to be a last resort.

Table 1.1. Common DNA lesions and the mechanism used for their repair.

Lesion	Effect on DNA	Causal agent	Repair mechanism
Cyclobutane pyrimidine dimers e.g. $\hat{T}T$ dimers.	Cross linking of neighbouring bases. Distorts the DNA helix and inhibits replication.	UV light	Photoreactivation, NER and BER (in prokaryotes and phage).
(6-4) photoproducts	Linkage of C6 and C4 of adjacent pyrimidines. Distorts DNA helix and inhibits replication.	UV light	Photoreactivation, NER and BER (in prokaryotes and phage).
Loss of purine or pyrimidine bases	AP site	Hydrolysis of N-glycosidic bond.	BER or NER.
Deamination e.g. Cytosine to Uracil.	Base changes, cause miscoding during replication	Spontaneous or nitrous acid exposure.	BER or NER.
Oxidised bases e.g. 8-hydroxyguanine.	Base changes, affects replication	Caused by exposure to ROS.	BER or NER.
Mis-incorporation of bases.	Base changes, can cause transitions and transversions.	Replication errors, DNA polymerases can insert the wrong base.	MMR
Alkylation of bases e.g. Adenine to $N^3$ -methyadenine.	Addition of methyl or ethyl groups.	Alkylating agents e.g. Nitrogen mustard	“The adaptive response” (Requires Ada and AlkB proteins) or NER.
DNA cross-links	Covalent linkage of DNA strands.	Alkylating agents e.g. Cis-platin and mitomycin C	Requires NER and HR.
Single-strand breaks	Gap in one strand of the deoxyribose sugar chains, affects replication	Ionising radiation and failure to join Okazaki fragments.	Repair by ligation or larger gaps by DNA polymerase.
Double-strand breaks	Both deoxyribose sugar chains are broken.	IR, bleomycin, endonucleases replication fork collapse	NHEJ or HR.



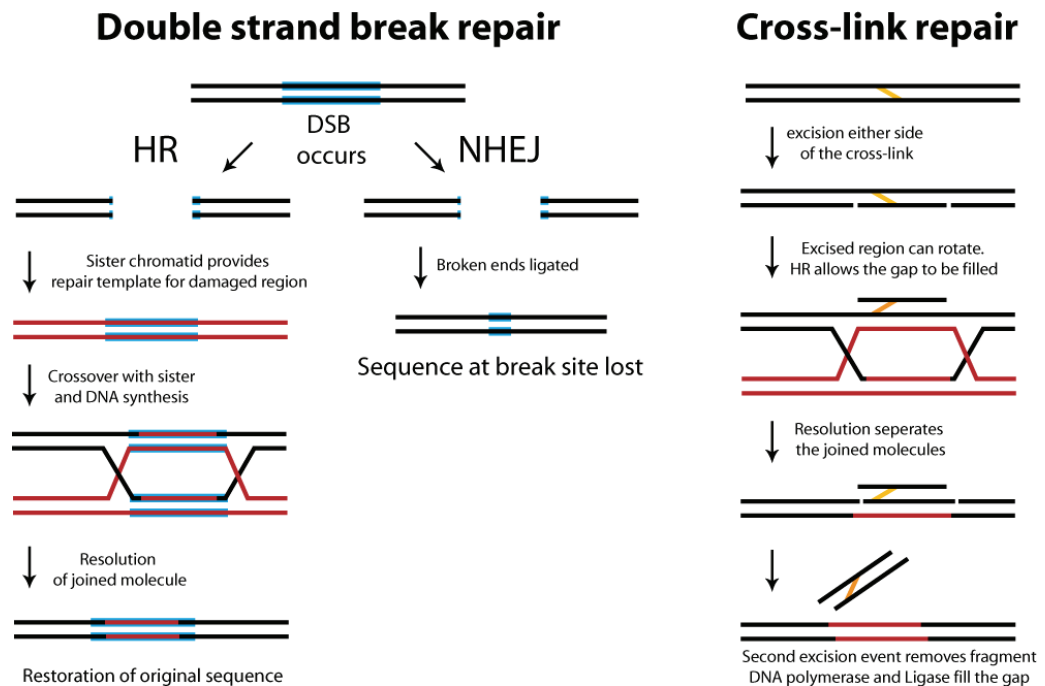


**Figure 1.1. Diagram outlining some of the excision repair pathways that organisms use to remove DNA lesions and preserve genomic integrity.**

The solid lines represent the deoxyribose sugar backbone while the small squares represent base pairs. Lesions and erroneous bases are represented in red, newly synthesised DNA is shown in black. BER involves a nitrogenous base being ‘flipped out’ of the DNA to create an apurine or apyrimidine (AP) site, which can be recognized by an AP endonuclease that breaks the sugar phosphate backbone removing the lesion. NER requires recognition of the distortion caused by a lesion, this allows recruitment of nucleases that nick either side of the lesion and a helicase, which removes the nicked fragment as a short oligonucleotide. The fragment liberated is always the same size: 12-13 bases in *E. coli* and 29 bases in mammals. MMR requires recognition of the DNA strands and of a mismatched base. The DNA is then nicked, either 3' or 5' of the lesion. A fragment containing the lesion is then removed; unlike in NER, this is done by degradation of the DNA by an exonuclease, assisted by a DNA helicase. Strand recognition is achieved by methylation state of the DNA (DNA methylation is shown as yellow methyl groups).

Lesions that affect both strands of the duplex, such as DSBs and DNA cross-links, are extremely genotoxic because they expose the DNA to degradation and/or block replication. DSBs can be repaired via one of two principal mechanisms, non-homologous end joining (NHEJ) or homologous recombination (HR) (Figure 1.2). The NHEJ pathway is often error-prone as during NHEJ the broken ends are ligated together in a reaction that may involve processing of the ends. This results in the loss of genetic material. The HR pathway requires homologous sequences to use as a template for the repair of the damaged region. This means that any sequences lost during the processing of the break can be copied from the template, preserving the DNA sequences around the break site (Figure 1.2) (Szostak *et al.* 1983; Critchlow

and Jackson 1998). DSBs are discussed in detail in Section 1.3 (cause) and 1.4 (repair). Inter-strand cross-links covalently join two strands of DNA. Their repair depends upon both NER and HR. It is thought that the cross-link is excised on one strand and HR then allows homologous sequences to be used to close the gap. NER can then remove the cross-linked region on the other strand and the duplex is restored by DNA synthesis (Cole 1973; McHugh *et al.* 2001).



**Figure 1.2. Diagram outlining the repair of some lesions, which affect both strands of the DNA duplex.**

The damaged DNA molecule is shown in black and its homologous sister, which maybe required for repair, is in red. The light blue background represents a region, which can be lost during NHEJ, but not HR. Cross-link repair involves both NER and recombination (covalent cross-links are indicated as a yellow line).

The range of lesions DNA can suffer is very extensive. Those discussed here and in Table 1.1 are just indicators of how varied the damage can be. Repair involves numerous, often overlapping, pathways and may require the cooperation of several systems to ensure removal of a lesion (Friedberg *et al.* 1995; Friedberg 2008). This is illustrated clearly by the systems evolved to

cope with oxidative damage caused by ROS, the most common endogenous source of damage to DNA.

## **1.2 Oxidative damage, damage prevention and repair**

The generation of ROS is an unavoidable consequence of respiration and the term oxidative stress has been coined to refer to the steady state of oxidative damage that occurs within cells. Oxidative stress has been implicated as a major contributor to ageing and human diseases. Experiments with murine models suggest that a calorie-restricted diet, which reduces the oxidative phosphorylation rates, can greatly reduce oxidative damage and prolong life (Aruoma *et al.* 1991; Sohal 2002). Most organisms have therefore evolved systems to cope with ROS. This includes pathways to detect and eliminate ROS before they can damage cellular components, including DNA. These mechanisms have been extensively studied in bacteria.

Two regulons have been identified in *E. coli* as having roles in the detection and elimination of ROS. Each comprises an oxidative stress sensor, a transcriptional activator and a set of enzymes concerned with the processing of ROS. The SoxRS regulon has the SoxR sensor protein and the SoxS transcription factor that activates genes in the regulon. The SoxRS system removes the superoxide anion. The *sod* genes, which encode superoxide dismutase (SOD), are major components of the SoxRS regulon. SOD converts the superoxide anion into H<sub>2</sub>O<sub>2</sub>, which, although less toxic can damage cellular components in micromolar concentrations. Therefore, H<sub>2</sub>O<sub>2</sub> also has to be detoxified and this is achieved by a second ROS detoxifying regulon, controlled by OxyR (Storz and Imlay 1999; Seaver and Imlay 2001b; Seaver and Imlay 2001a; Seaver and Imlay 2004).

The OxyR protein is a dimer that acts both as a sensor of oxidative stress and as the transcriptional activator for the regulon. As generation of H<sub>2</sub>O<sub>2</sub> is unavoidable, the OxyR regulon has an important housekeeping function (Seaver and Imlay 2001b; Seaver and Imlay 2004). It is found in a large number of bacteria and has been shown to be vital for survival after exposure to ROS, including H<sub>2</sub>O<sub>2</sub>, which are often encountered when a

bacterial pathogen tries to invade a host. The bacteria are attacked by the host's immune cells, which utilise an oxidative burst of  $\text{H}_2\text{O}_2$  to kill the pathogen. This makes OxyR an important virulence factor (Mukhopadhyay and Schellhorn 1997; Ochsner *et al.* 2000).

As oxidative stress increases so does the redox potential inside a cell. This permits the formation of a disulphide bond between the two subunits of the OxyR dimer. The conformational change this induces allows the protein to bind to its target sequences and activate expression of proteins in the regulon (Zheng *et al.* 1998). The OxyR protein induces several genes, including itself (Volkert and Landini 2001) and the catalase genes, *katG* and *katH*, which encode two separate catalases, both capable of breaking down  $\text{H}_2\text{O}_2$  (Mukhopadhyay and Schellhorn 1997). However, the AhpC protein, rather than the catalases, has been identified as the primary  $\text{H}_2\text{O}_2$  scavenger in *E. coli* (Seaver and Imlay 2001a).

AhpC is part of the OxyR regulon, along with its co-factor AhpF. Together they are able to break down  $\text{H}_2\text{O}_2$  (Storz *et al.* 1989; Tartaglia *et al.* 1989). This reaction differs from the catalase-driven breakdown in its kinetics. AhpC can degrade  $\text{H}_2\text{O}_2$  at very low concentrations ( $<0.1\mu\text{M}$ ), whereas catalases require  $\text{H}_2\text{O}_2$  concentrations above  $1\mu\text{M}$  (Seaver and Imlay 2001b). This means AhpC can prevent ROS build up and thus limit DNA damage. Deletion of both the catalase genes does not affect  $\text{H}_2\text{O}_2$  levels within the cell. A detectable increase in  $\text{H}_2\text{O}_2$  levels requires a triple deletion of both catalase genes and the *ahpC* gene (Seaver and Imlay 2001a; Seaver and Imlay 2001b). AhpC and AhpF are both required for the removal of  $\text{H}_2\text{O}_2$ . AhpC degrades  $\text{H}_2\text{O}_2$  but then has to be regenerated by AhpF before it can remove any more  $\text{H}_2\text{O}_2$ . The regeneration is energy intensive, requiring NADH (Poole and Ellis 1996; Poole 2005). In contrast, catalases do not require regeneration, so although AhpC is a highly efficient  $\text{H}_2\text{O}_2$  detoxifying enzyme it can become saturated at moderate  $\text{H}_2\text{O}_2$  concentrations and its activity is energy intensive.

Evidence from several studies suggests interplay between the ROS scavenging systems and a cell's DNA repair pathways. Expression of bacterial AhpC in human cells reduces nuclear DNA damage and inhibits apoptosis

(Lombard *et al.* 2005). Furthermore, OxyR is involved in the upregulation of *uvrD*, which encodes a DNA helicase, known to be involved in several DNA repair pathways, including NER and MMR (Figure 1.1) (Modrich 1994; Mukhopadhyay and Schellhorn 1997). Studies of bacterial responses to oxidative stress found that several genes involved in HR are important for resistance to H<sub>2</sub>O<sub>2</sub>. Indeed, increased sensitivity to H<sub>2</sub>O<sub>2</sub> was observed in both *recA* and *recBCD* *E. coli* mutants, which is probably a combined result of these mutants being deficient in both HR and induction of the SOS response, which provides a global response to DNA damage (Linn and Imlay 1987; Asad *et al.* 1997; Konola *et al.* 2000; Erill *et al.* 2007). In *Neisseria gonorrhoea*, *recA* and the recombination associated gene *recN*, are both important for survival after H<sub>2</sub>O<sub>2</sub> exposure and *recN* is also important for H<sub>2</sub>O<sub>2</sub> resistance in *Helicobacter pylori* where it is induced by H<sub>2</sub>O<sub>2</sub> exposure (Stohl and Seifert 2006; Wang and Maier 2008). In *Pseudomonas aeruginosa*, *oxyR* is in an operon with *recG*, which is known to be involved in DNA repair (Ochsner *et al.* 2000). Taken together these data suggest a degree of co-regulation of the DNA repair and the oxidative stress responses. This gives a layered defence against oxidative stress and the damage it causes, to ensure the preservation of genomic integrity against this constant threat.

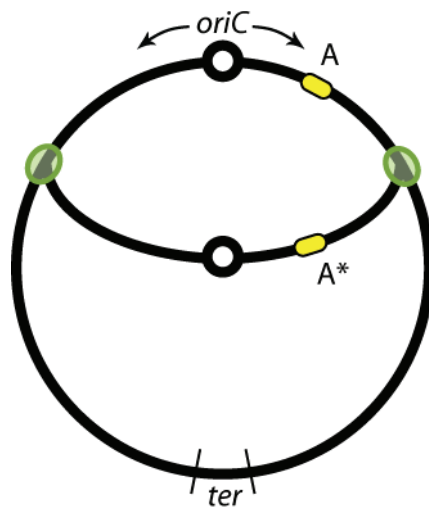
### 1.3 DSB formation

A DNA DSB is probably the most genotoxic lesion a cell can encounter (van Gent *et al.* 2001). Not only are both strands of the DNA backbone broken, but spatial separation of the broken ends can also occur. Thus, repair of a DSB is not only essential, but complicated by the need to keep the broken ends in proximity to each other. DSBs can be classified according to whether they are generated in a replication-independent or replication-dependent manner and whether one or two DNA ends are exposed.

Generally, replication-independent breaks are caused by direct attacks on the DNA and result in two exposed DNA ends (Figure 1.2). The lethality of IR exposure is largely due to the DNA DSBs it causes; a dose of just 1 Gray can cause 40 DSBs in a mammalian cell (Ward 1988). Exactly how IR causes a

break is unknown. Both direct cleavage of the DNA and clustered damage caused by ROS, generated by radiolysis of water in the cell, have been suggested (Siddiqi and Bothe 1987). High UV doses can also cause replication-independent breaks, which are thought to arise as a result of excision repair of damaged bases on both DNA strands at nearby sites. Chemical agents, such as the anti-cancer drug bleomycin and the antibiotic mitomycin C, have also been shown to cause DSBs (Povirk *et al.* 1977), although mitomycin C is actually a DNA cross-linking agent and thus is not a direct cause of DSBs (Iyer and Szybalski 1963). Cellular endonucleases can also attack DNA, causing a DSB. Ironically, eukaryotes utilise endonucleases to deliberately break a chromosome and thus induce meiotic recombination (Keeney *et al.* 1997). Likewise, the Rag1 and Rag2 enzymes create DSBs to trigger V(D)J recombination, which is required to generate the diversity seen in antibodies. Yeast makes use of site-specific HO endonucleases to create site-specific DSBs that trigger recombination and facilitate mating type switching (Haber 1995; Belfort and Roberts 1997; Keeney *et al.* 1997). All of the replication-independent systems will generate a DSB like that in Figure 1.2, where two DNA ends have been generated.

During normal growth the vast majority of DSBs are likely to arise in a replication-dependent manner. DNA replication has evolved to be highly accurate and processive and has been extensively studied in *E. coli*. The *E. coli* chromosome consists of a single circular DNA molecule, with a single origin of replication called *oriC*. Replication occurs in a bi-directional manner from the origin, with two replication forks being established at the origin, which then move around the chromosome in opposite directions. Approximately 180° around the chromosome from *oriC* the forks encounter the termination region (*ter*) where replication is completed. The two chromosomes are then separated and segregated into daughter cells (Figure 1.3).



**Figure 1.3. Schematic representation of a replicating *E. coli* chromosome.**

The replisomes (green circles) and DNA (in black) are shown. If a DSB were to occur at the site marked A then there is an intact copy within the newly replicated regions of DNA at A\*, which can be used as a template. The origin (*oriC*) and terminus (*ter*) regions are marked.

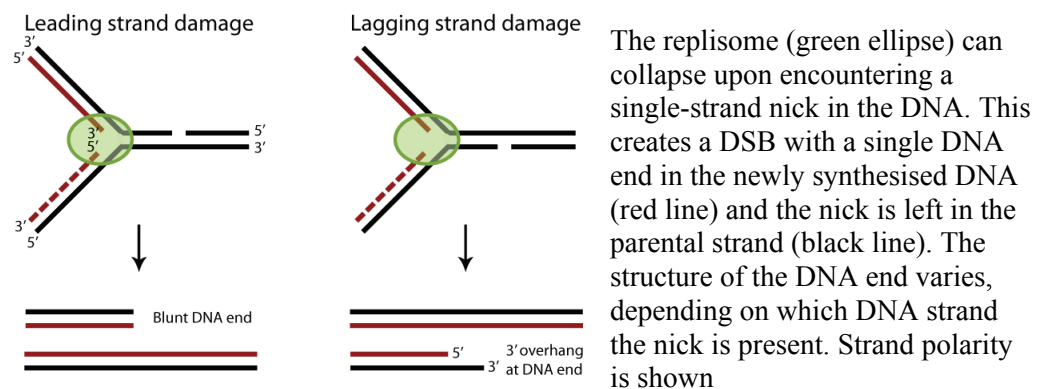
The leading DNA strand is replicated continuously, while the lagging strand is synthesized as short fragments, of approximately 1000 nucleotides, called Okazaki fragments, which are subsequently ligated together (Kornberg and Baker 1992). Complete replication of the chromosome takes approximately 40 minutes. However, exponentially growing *E. coli* can divide every 20 minutes. This means that there must be multiple, overlapping rounds of replication and as a result there can be multiple copies of at least portions of the chromosome. Consequently if a DSB was to occur in the replicated region, then there would also be an intact copy of this DNA, which can be used as a repair template (Figure 1.3).

Despite accurate and rapid replication being vital, replication forks are commonly perturbed before they get to the terminus. This is often due to DNA lesions, although secondary structure and DNA bound proteins also have an effect (Hyrien 2000). Estimates in bacteria suggest a replication fork rarely completes its journey around the chromosome without encountering some form of blockage, which can lead to the formation of a DSB (Cox *et al.* 2000; Sandler and Marians 2000; Marians 2004).

The majority of replication-dependent breaks arise when a replication fork collapses due to passage through a single-strand nick (reviewed by Kuzminov 1995b). Nicks can be generated on the lagging strand as a result of the removal of the RNA primers that facilitate the semi-discontinuous

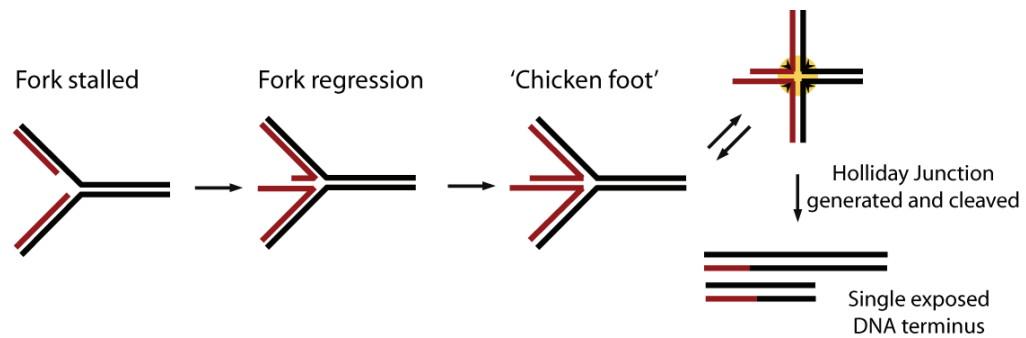
replication of this strand. However these will tend to be sealed by DNA ligase (Heitman *et al.* 1989). Nicks that arise due to DNA damaging agents will occur with equal frequency on both strands. When the replication fork encounters a nick, it can cause fork collapse, thus generating a DSB with a single exposed end (Figure 1.4) (Cox *et al.* 2000; Sandler and Mariani 2000; Kuzminov 2001). A noteworthy feature of replication-dependent breaks arising from a nicked leading strand is that they are blunt ended, whereas if they arise as a result of a nicked lagging strand they are likely to have a 3' overhang (see Figure 1.4).

**Figure 1.4. Diagram describing the formation of replication-dependent DSBs.**



The collapse of the replisome is not pre-determined by an encounter with a replication block. Sometimes the replisome can stall, but remain assembled and bound to the DNA at the site of a blockage (McGlynn and Lloyd 2000). To allow replication to proceed the cause of the block must be removed. In the case of a DNA lesion this requires repair. However, the replisome prevents repair proteins gaining access to the lesion and must therefore, be removed to allow repair. This is achieved by fork regression, which involves annealing of the newly synthesised DNA to create a type of Holliday junction, referred to as a 'chicken foot'. Holliday junctions can be recognised and cleaved by structure specific endonucleases, resulting in a DSB (Michel *et al.* 1997; Seigneur *et al.* 1998; McGlynn and Lloyd 2000; Seigneur *et al.* 2000; Postow *et al.* 2001) (Figure 1.5).





**Figure 1.5. Generation of DSBs from a stalled replication fork.**

Fork regression can be driven in several ways, but leads to the generation of a 'chicken foot' type Holliday junction. This can be cleaved, resulting in the generation of a DSB, with one exposed terminus.

Generally, replication-independent breaks result in two exposed DNA ends, while the replication-dependent breaks result in one. However, replication-dependent breaks with two exposed DNA ends can also arise. The SbcCD complex of *E. coli* is a specialized nuclease capable of cleaving hairpin structures. These are proposed to arise during the replication of palindromic sequences in the lagging strand template, which is transiently single-stranded. The DNA opposite the hairpin cannot be synthesised and subsequent cleavage of the hairpin by SbcCD generates a DSB (Connelly *et al.* 1999; Cromie *et al.* 2000; Eykelenboom *et al.* 2008). DSBs can even arise because of MMR (Figure 1.1). In cells deficient for the *dam* gene product, which methylates DNA, the parental and newly synthesised DNA strands cannot be distinguished and excision repair can initiate on both, which results in a DSB if the excision events overlap (Au *et al.* 1992).

Studies on the repair of DSBs have been conducted in numerous organisms including mammals, yeast and bacteria. These studies often made use of IR and chemical mutagens to generate DSBs. However, such agents cause a wide variety of DNA lesions and other cellular damage. The spectrum of damage caused makes it impossible to assign the observed phenotypes to the presence of DNA DSBs alone. To avoid this uncertainty, many studies have focused on analysing specialised recombination events, which are often initiated by the cell inducing a DSB, such as V(D)J recombination of

mammalian immunoglobulin genes and the mating type switching of yeast (Haber 1998b; Pierce *et al.* 2001; Sugawara and Haber 2006). Studies in *E. coli* have focused on analysis of the phenotypes observed in wild-type and various mutant strains after exposure to IR, restriction enzyme attack, palindrome cleavage or recombination events associated with linear DNA, like conjugation and transduction (Sargentini and Smith 1986; Asai *et al.* 1994; Lloyd and Low 1996; Cromie *et al.* 2000; Cromie and Leach 2001). To overcome these problems, systems have been developed to generate site specific DSBs using the yeast homing endonuclease, I-SceI. These systems allow defined DSBs to be generated at particular loci within the genome and have been used successfully in mammalian and yeast cells (Lukacsovich *et al.* 1994; Haber 2000; Johnson and Jasin 2001). However, only recently were similar systems developed for bacteria, where an inducible I-SceI endonuclease has been used to cleave the *E. coli* chromosome at engineered target sites (Chapter 4) (Meddows 2002; Meddows *et al.* 2004; Grove *et al.* 2008). The specificity of these systems allows far more conclusive analysis of consequences of DSB formation and of the repair of DSBs *in vivo*.

## 1.4 DSB repair

The integrity of the DNA backbone is constantly threatened, making the occurrence of DSBs endemic. Cells have therefore evolved systems to cope with DSBs, repairing them so as to preserve the integrity of the genome (Paques and Haber 1999; Cox 2001). There are two distinct DSB repair pathways, which have been identified in eukaryotic cells, HR and NHEJ (Figure 1.2) (Hoeijmakers 2001; van Gent *et al.* 2001). HR can accurately repair a DSB, but requires an intact copy of the damaged region to act as a template. In contrast NHEJ repairs DSBs without a template, but the process is potentially less accurate.

NHEJ involves numerous proteins that capture the broken ends and then process them so that they can be ligated. This process can lead to loss of nucleotides, both during the breakage and subsequent processing, posing a threat to genomic integrity. After a DSB occurs, the exposed DNA termini can

be bound by heterodimers of the Ku70 and Ku80 proteins. The Ku-dimer has an open ring-shaped structure and can load onto exposed DNA ends (Walker *et al.* 2001; Mari *et al.* 2006; Uematsu *et al.* 2007). The Ku-dimers act as a scaffold, recruiting the DNA-dependent protein kinase catalytic subunits (DNA-PKcs) to the DNA ends. This kinase has several roles in DNA repair, but most importantly perhaps, it can interact with the DNA-PKcs binding the other end of a DSB to form a synaptic complex. The complex bridges the two DNA ends and holds them stably together, so they can be ligated (Yaneva *et al.* 1997; DeFazio *et al.* 2002).

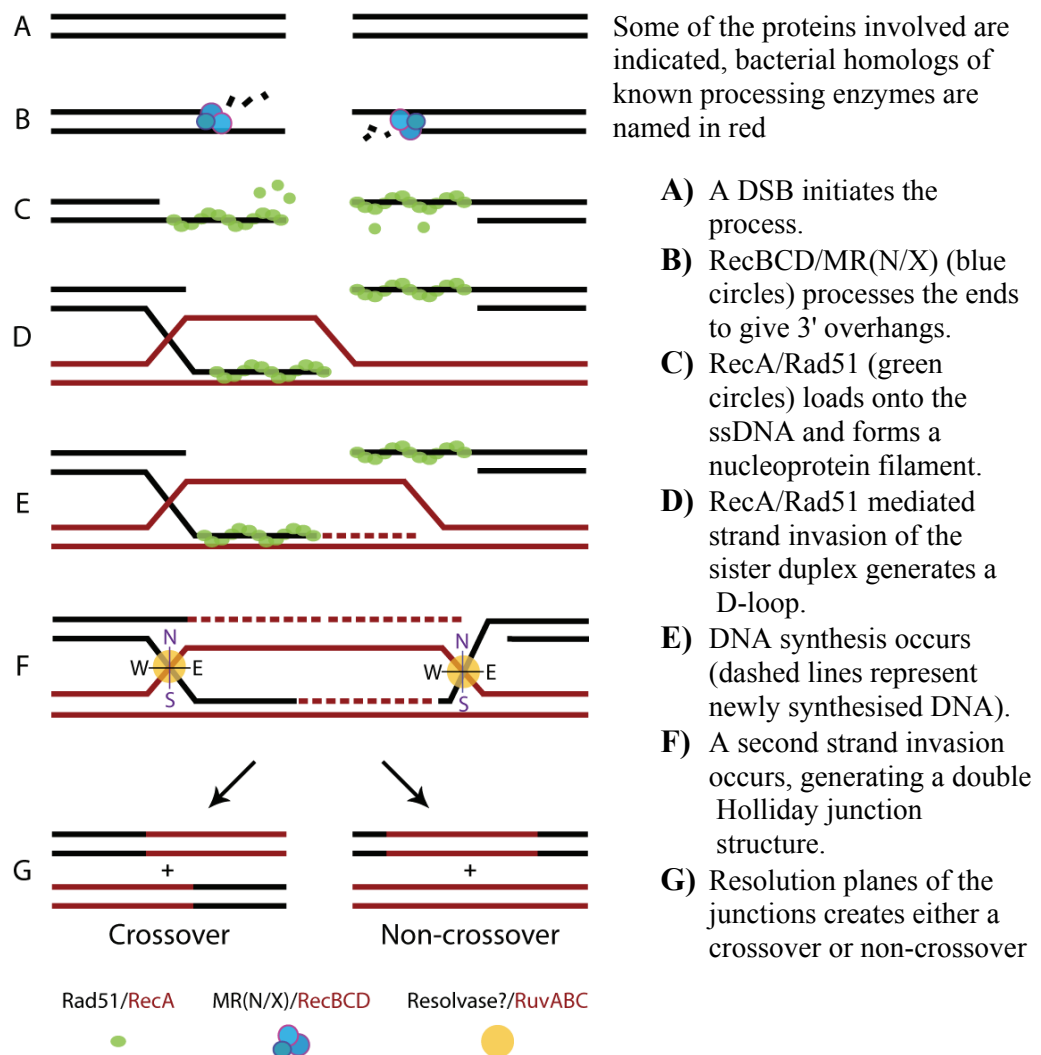
Ligation requires DNA ends to be blunt, therefore overhangs have either to be filled in by polymerases, or resected by a nuclease, specifically Artemis, but possibly the MR(N/X) complex (Jeggo *et al.* 1998). DNA polymerase  $\mu$  has been shown to interact with Ku proteins (Mahajan *et al.* 2002) and depletion of DNA polymerase  $\lambda$  reduces end joining *in vitro* (Lee *et al.* 2004). Artemis has a crucial role in V(D)J recombination and has been demonstrated to interact with DNA-PKcs, while mutants are sensitive to IR (Ma *et al.* 2002).

The ligation reaction restores the DNA strands and requires the action of a protein complex comprising XRCC4 and DNA ligase IV. Exactly when this complex is recruited to a break is unclear, but it does interact with Ku, suggesting it could be recruited early during repair (Nick McElhinny *et al.* 2000; Mari *et al.* 2006; Costantini *et al.* 2007). It also seems that the interaction with the DNA-Ku scaffold enhances the efficiency of ligation (Nick McElhinny *et al.* 2000).

The second pathway of DSB repair is HR. The classical model was first proposed by Szostack (1983) to explain meiotic recombination in yeast (Figure 1.6). In this model a DSB is processed via a four way branched structure, referred to as Holliday Junction. These structures were first proposed by Robin Holliday in 1964 as part of his model to explain recombination in fungi (Holliday 1964). However, unlike the original Holliday model, the Szostack model predicts an intermediate structure containing two Holliday junctions (Holliday 1964; Szostak *et al.* 1983). HR in eukaryotes was first studied

extensively in the ascomycete fungi and budding yeast *Saccharomyces cerevisiae*. Many of the yeast genes involved were identified as mutations leading to radiation sensitivity and were thus named RAD genes, including *RAD50*, *RAD51*, *RAD52* etc. (Krogh and Symington 2004). Many of these genes are conserved between yeast and man, suggesting an evolutionary conservation of HR processes. HR begins with the processing of the DNA ends at a DSB, to allow synapsis of two homologous pieces of DNA. Synapsis requires strand invasion of an intact homologous duplex, which is achieved by the action of three proteins RAD51, RAD52 and RPA (McIlwraith *et al.* 2000). It is proposed that the exposed termini at a break site (Figure 1.6 A) are processed to generate 3' single-stranded DNA (ssDNA) overhangs or tails (Figure 1.6 B), by the exonuclease activity of the MRE11/RAD50/NBS1 complex (XRS2, replaces NBS1 in yeast), which is also involved in the early binding and recognition of DSBs (Krogh and Symington 2004). The overhangs are bound by a recombinase, RAD51 (a homolog of bacterial RecA), which forms a nucleoprotein filament and guides the invasion of the homologous duplex DNA (Figure 1.6 C) (Paques and Haber 1999; Allers and Lichten 2001b). RAD52 actively recruits RAD51 to the ssDNA and facilitates its loading by displacement of the single-strand binding protein RPA, from the tails (Shinohara *et al.* 1992; Sugiyama and Kowalczykowski 2002; Sugiyama *et al.* 2006). Strand invasion generates a specific structure called a D-loop (Figure 1.6 D) in which the invading single-stranded tail can act as a primer for DNA synthesis (Figure 1.6 E).

**Figure 1.6. Model of DSBR adapted from Szostak (1983).**



Repair may then proceed via the classical DSB repair model as described by Szostak (1983), with the other exposed DNA end participating in a second strand invasion event generating a double Holliday junction structure (Figure 1.6 F) (Sugiyama *et al.* 2006). Specialised proteins, termed resolvases, can break the junctions formed. These resolvases are well characterised in bacteria, but remain elusive in eukaryotes. In bacteria, resolution is often coupled to ‘branch migration’ of the Holliday junctions, moving the junction so as to expand the region of homology and also allows the resolvase to locate its target sequence and cleave the DNA. Resolution generates either a crossover or a non-crossover product, depending on which plane the two Holliday junctions are resolved. (Figure 1.6 G).

As an alternative, synthesis-dependent strand annealing (SDSA) could occur. Based on yeast studies, it is believed that the D-loop generated during strand invasion primes replication (Figures 1.6 D and 1.7 A), after which the newly synthesised DNA can be displaced and annealed to its broken partner (Figure 1.7 B). Repair can then be completed by replication of the missing DNA as if it were a gap in the duplex (Figure 1.7 C). In this model SDSA generates non-crossovers, while crossovers arise from HR (Allers and Lichten 2001a). DNA synthesis, primed from a D-loop, is essential in both pathways (Allers and Lichten 2001a; McIlwraith *et al.* 2005; McIlwraith and West 2008). Several repair polymerases have been suggested to facilitate DNA synthesis from a D-loop (Goodman 2002), but only DNA polymerase  $\eta$  has an identified role in HR. In the absence of polymerase  $\eta$ , the rate of gene conversion is reduced in chicken DT40 cells, furthermore, it has been demonstrated to synthesise DNA from a D-loop (Figure 1.6 D) (McIlwraith *et al.* 2005; McIlwraith and West 2008).



**Figure 1.7. SDSA pathway of DSB repair as proposed by Allers and Lichten (2001).**

- A)** Strand invasion, guided by Rad51/RecA (green circles), generates a D-loop allowing DNA synthesis to be primed by PriA.
- B)** Newly synthesised DNA (dashed line) is displaced and annealed to its sister strand.
- C)** Missing DNA is synthesised as if it were a gap in the DNA, generating a non-crossover event. Crossovers are generated by HR.

Both DSB repair pathways share a common start point and possibly the same initial detection and processing mechanisms (Aylon and Kupiec 2004). It is unclear what controls the choice of repair pathway and pathway preference may vary depending on the cell cycle. HR is an option, only when an intact

homologous sequence is available to act as a template for DNA synthesis (Shrivastav *et al.* 2008). Mutations in BRCA1 and BRCA2 restrict HR, making NHEJ the only pathway for the repair of DSBs (Moynahan *et al.* 1999; Moynahan *et al.* 2001). This leads to major genomic instability, including chromosome translocation and deletions, and is suggested to be the cause of the cancer predisposition associated with BRCA1 and BRCA2 defects (Yu *et al.* 2000; Wang *et al.* 2001). It seems that both HR and NHEJ have key roles in DSB repair, despite the mutagenic nature of NHEJ (Stephanou *et al.* 2007).

The two models discussed involve the processing of both DNA ends. However, HR can involve just one exposed DNA end, either as a result of only one successfully invading, or it could be a replication-dependent break, with only one end. It has been thought for some time that one-ended recombination events could lead to the establishment of a replication fork (Meselson and Weigle 1961). This idea was developed by Kogoma to explain how replication in *E. coli* could be initiated at sites other than *oriC* and he proposed a process of break-induced replication (BIR) (Kogoma 1996; Kogoma 1997). In eukaryotes, BIR is thought to allow non-reciprocal recombination, which can be induced in yeast by a DSB and is dependent on RAD51 and RAD52 (Malkova *et al.* 1996; Bosco and Haber 1998). Data suggests that BIR is important for rescuing broken replication forks (Kuzminov 1995b; Seigneur *et al.* 1998; Michel 2000; Kraus *et al.* 2001; Michel *et al.* 2001) and involves the formation of a D-loop, although whether a full replication fork is then established is currently unclear (Kraus *et al.* 2001).

#### **1.4.1 DSB repair in bacteria**

Until recently, it had been assumed that prokaryotes repaired DSBs solely by HR. However, NHEJ has now been identified in several bacterial species, initially in *B. subtilis* and then several species of mycobacteria where Ku-like proteins and a specific end-joining ligase, LigD have been identified (Doherty *et al.* 2001; Weller *et al.* 2002; Gong *et al.* 2005; Shuman and Glickman 2007).

Ironically, despite most of the DSBR models being based on studies of eukaryotic systems, in particular yeast, the enzymology of DSB repair via HR is probably best understood in *E. coli*. The recombination events initiated at

DNA ends are particularly well studied. The exposed DNA termini are usually processed by RecBCD to facilitate recombination dependent DNA repair (Spies *et al.* 2005). The RecBCD complex is comprised of three subunits. The RecB and RecD subunits are helicases, with the opposing polarities. This allows them to load onto different DNA strands of the duplex, yet translocate overall in the same direction. RecC has regulatory roles, controlling the activity of the complex (Boehmer and Emmerson 1991; Yu *et al.* 1998; Dillingham *et al.* 2003; Taylor and Smith 2003). To initiate repair, RecBCD binds to blunt DNA termini at the break site. It then unwinds the duplex with high processivity and degrades both DNA strands (Roman *et al.* 1992). As RecBCD translocates, the RecC subunit can scan for  $\chi$  sequences (chi- crossover hotspot instigator, 5'-GCTGGTGG-3') (Lam *et al.* 1974). Upon encountering  $\chi$ , RecC binds to the exposed 3' strand and attenuates the complex's nuclease activity, resulting in the generation of 3' ssDNA tails, with  $\chi$  sequences at their ends (Figure 1.6 B). By this action RecBCD expands a break into a gap of over 5Kb on average (Taylor *et al.* 1985; Taylor and Smith 1995; Bianco and Kowalczykowski 1997; Handa *et al.* 1997; Singleton *et al.* 2004).

RecBCD can then facilitate loading of RecA onto the single-stranded tail, forming a nucleoprotein filament (Figure 1.6 C) (Anderson and Kowalczykowski 1997). This filament subsequently acts to promote homologous pairing and strand exchange of DNA, during which the ssDNA invades an intact sister duplex, displacing one of the DNA strands to generate a D-loop (Figure 1.8 D) (Rao and Radding 1993; Anderson and Kowalczykowski 1997; Churchill and Kowalczykowski 2000; Galletto *et al.* 2006; Spies and Kowalczykowski 2006). D-loops are believed to allow priming of DNA replication, via the action of PriA protein (McGlynn *et al.* 1997) and DNA synthesis is required to fill in the missing DNA at the break site (Figure 1.6 E) (Liu *et al.* 1999). Expansion of the region paired by RecA leads to reciprocal exchange of the DNA and can create a Holliday junction. This leads to the formation of a replication fork, associated with a Holliday junction (as seen in Figure 1.8). If the second tail is processed in the same manner, then instead of the double Holliday junction structure envisaged by



Szostack *et al* (1983) (Figure 1.6 F), two replication forks, each associated with a Holliday junction, are established some 5kb or more apart.

In order to complete the repair process, branch migration and cleavage of the Holliday junctions is necessary to separate the synapsed duplexes. These reactions are achieved in *E. coli* by the RuvABC complex. A tetramer of RuvA forms a platform on which the Holliday junction can be held in an open square planar conformation. This is believed to involve four protrusions at the centre of the tetramer, each made up of a pair of acidic residues, which form acidic ‘pins’ that separate the DNA strands (Parsons *et al.* 1995; Rafferty *et al.* 1996). RuvB is a hexameric ring helicase that assembles on each of two diametrically opposed arms of a Holliday junction, bound by RuvA (Mitchell and West 1994; Stasiak *et al.* 1994). Branch migration is ATP dependent and is achieved by both RuvB rings pulling DNA through the complex across the channels on the surface of the RuvA tetramer (Tsaneva *et al.* 1992; Parsons *et al.* 1995). RuvC is proposed to associate directly with RuvAB and the complete RuvABC complex has been termed the “resolvasome” (Kuzminov 1993; Mandal *et al.* 1993; Whitby *et al.* 1996; Davies and West 1998; van Gool *et al.* 1998). RuvC is a dimeric, structure specific endonuclease that can scan for its consensus sequence (5'-<sup>A</sup>/<sub>T</sub>TT↓<sup>G</sup>/<sub>C</sub>-3', where ↓ indicates nick site) as the DNA is pulled across RuvA. When the sequence is detected, the RuvC dimer nicks opposing DNA strands of the junction (Figure 1.6 F) (Bennett *et al.* 1993; Shah *et al.* 1997; Davies and West 1998). Nicking by the resolvase can occur in one of two planes, to give either a crossover or a non-crossover event, which should occur in the same ratio (Figure 1.6 F) (Davies and West 1998; van Gool *et al.* 1999).

The original evidence that both the branch migration and resolution reactions are linked to the same enzyme complex came from genetic studies. As mutations in any of the *ruv* genes gave the same phenotype and mutations of any of the three could be suppressed by expression of an alternative resolvase, RusA (Mandal *et al.* 1993). RusA is normally silent, present in *E. coli* on a cryptic prophage, DLP12 (Mandal *et al.* 1993; Mahdi *et al.* 1996). When expressed, RusA will cleave branched DNA structures, particularly

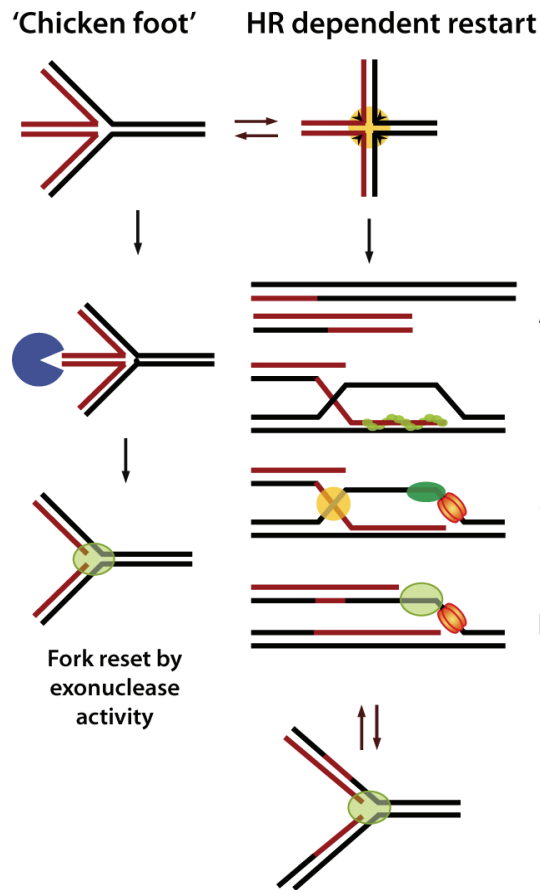
Holliday junctions, although it is less structure specific than RuvC (Sharples *et al.* 1994; Chan *et al.* 1997; Giraud-Panis and Lilley 1998). However, the ability of RusA to suppress *ruv* mutants depends on the activity of another protein, RecG (Mandal *et al.* 1993). RecG is a highly conserved bacterial helicase, with the ability to branch migrate Holliday junctions and other branched DNAs. As long as RecG is present, *ruv* mutants remain recombination proficient, even though they are defective in DNA repair suggesting RecG can provide a parallel pathway for some of the Ruv proteins functions (Lloyd 1991; Lloyd and Sharples 1993a; Lloyd and Sharples 1993c; Sharples *et al.* 1999; Wen *et al.* 2005).

RecG also appears to have an important role in the processing of stalled replication forks, being capable of regressing them to create the ‘chicken foot’ structure (McGlynn 2000; McGlynn and Lloyd 2001; Gregg *et al.* 2002; Wen *et al.* 2005). A reversed fork could in theory be restarted by nuclease processing of the ‘chicken foot’. Alternatively, a recombination-based pathway could restore the fork (Figure 1.8). This leads to the generation of a DSB with one exposed DNA end (Figure 1.8 A). This must be repaired and replication restarted (Figure 1.8 A – E). First, as in DSB repair, the exposed end is processed to generate a 3' ssDNA tail, onto which RecA can be loaded promoting strand invasion to generate a D-loop (Figure 1.8 B). This structure can be used to initiate replication and restore the chromosome.

In *E. coli* the initiation of replication at sites other than *oriC* requires the action of PriA. PriA has been known for some time to be involved in replication. *In vitro* it is required for the assembly of the  $\Phi$ X174 primosome (Wickner and Hurwitz 1975; Shlomai and Kornberg 1980). A *priA* mutant shows an array of phenotypes, including replication deficiencies, low viability and increased UV and IR sensitivity (Sandler and Marians 2000). Importantly it has been shown that PriA can bind and assemble a replisome at D-loop structures, by recruitment of the replicative helicase DnaB (Figure 1.8 C and D) (McGlynn *et al.* 1997; Liu and Marians 1999). In *E. coli* it seems likely that the majority of breaks would be processed in this manner. Even if a two-ended DSB occurs, the rampant DNA degradation caused by RecBCD would mean

that the ends were several kilobases apart and so be processed as if they were two separate single-ended breaks (Roman *et al.* 1992).

A reliance on recombination proteins to restart replication forks would also explain the relatively low viability of *recA* and *recB* mutants (approximately 50% and 75% viability, respectively) (Capaldo-Kimball and Barbour 1971). Mutations in *polA* and *lig* mutations, which encode the *E. coli* repair polymerase and DNA ligase respectively (Kornberg and Baker 1992; Kuzminov 1995b), prevent the maturation of Okazaki fragments into a complete DNA strand. This increases the occurrence of fork collapse and *polA* and *lig* mutations are synthetically lethal in conjunction with a *recA* mutation (Kuzminov 1995b), suggesting that without recombination, collapsed forks cannot be restarted. Since replication forks rarely reach the terminus unhindered, fork restart would appear to be a housekeeping function of the recombination machinery (Cox *et al.* 2000; Mariani 2000; Sandler and Mariani 2000; Mariani 2004).



**Figure 1.8. Potential pathways for the restart of a stalled fork.**

A regressed fork structure (see Section 1.3 and Figure 1.5 for details) could be processed directly by endonucleases to reset the fork or by a recombination-mediated pathway.

- A)** The recombination pathway involves the generation of a single-ended DSB.
- B)** The break is processed and invades a sister duplex, to form a D-loop.
- C)** This is bound by PriA (dark green ellipse), which recruits DnaB helicase (orange ellipses)
- D)** DnaB allows loading of the replisome.
- E)** The replisome (green ellipse) is reloaded. A resolvase (orange circle), maybe required to break a Holliday junction formed.

The repair of a DSB is usually initiated by the processing of the DNA ends by RecBCD. In its absence, recombination is barely detectable and cells are sensitised to agents that damage DNA. However, suppressor mutations can arise that restore almost wild-type levels of recombination and viability to cells. These suppressor mutations, termed suppressors of *recBC* (*sbc*) mutations, are found in two forms, both of which allow recombination to occur in a RecBCD-independent manner. The first class, *sbcA* mutations, occur only in *E. coli* strains carrying the Rac prophage. The prophage encodes two genes, *recE* (encoding exonuclease VIII) and *recT*, and expression of both is activated by *sbcA* mutations. It is proposed that RecE replaces RecBCD's nuclease activity, while RecT acts in place of RecA to promote strand exchange, as RecA loading is reduced in the absence of RecBCD (Kolodner *et al.* 1994; Lloyd and Low 1996). The second class of *sbc* suppression requires mutations in two unlinked genes, one in *sbcB*, and the second in either *sbcC* or *sbcD*. The

*sbcB* mutation affects exonuclease I, while *sbcC* and *sbcD* encode the two components of the heterodimeric nuclease SbcCD. As these suppressor mutations inactivate nucleases, DNA ends persist giving greater opportunity for RecA to be loaded (Lloyd and Low 1996). It is thought that in a *recBC sbcBC* background, end processing requires the action of RecQ helicase and RecJ nuclease (Kowalczykowski 2000), with loading of RecA facilitated by the action of three proteins; RecF, RecO and RecR. In wild-type cells, these proteins are responsible for directing RecA loading at ssDNA gaps but the RecFOR proteins can displace SSB and load RecA at DNA ends, albeit less efficiently than RecBCD (Kushner *et al.* 1971; Lloyd and Buckman 1985; Morimatsu and Kowalczykowski 2003). Once RecA is loaded, recombination could proceed as described above.

## 1.5 RecN

The *recN* gene was identified independently by two laboratories, in 1983. It was isolated by Lloyd *et al.* (1983) as a DNA damage inducible and recombination-defective Mud(Ap<sup>r</sup>) insertion mutation in a *recB sbcBC* strain background. Sargentini and Smith (1983) identified a randomly generated mutant, sensitive to IR, that they designated *radB101*. They later mapped *radB101* to the same location as *recN*. Since this was the more commonly used nomenclature, *recN* was retained (Sargentini and Smith 1988). Early studies suggested that RecN was involved in the repair of DSBs and regulated as part of the *E. coli* SOS response (Lloyd *et al.* 1983; Sargentini and Smith 1983; Picksley *et al.* 1984a; Picksley *et al.* 1985a).

### 1.5.1 The SOS response

Found in many bacteria, the SOS response provides a global response to DNA damage and has become the textbook paradigm of a coordinated gene response. It leads to the induction of numerous genes involved in the repair of DNA lesions, including members of the NER, HR and lesion bypass pathways. A DNA damage response in *E. coli* (reviewed by Witkin (1976)) was first described in the 1970s and following studies that suggested a common basis for many, apparently disparate observations, giving rise to the theory of a

global DNA damage response named the SOS response (Radman 1974; Radman 1975). The SOS regulon works via the repression of member genes by the action of LexA, which binds to a consensus TACTG(TA)<sub>5</sub>CAGTA motif, often referred to as an SOS box (Berg 1988). These are positioned to interfere with RNA polymerase binding or translocation along the DNA when LexA is bound. Variations of the SOS box sequences, away from the consensus sequence, means that LexA binds to different SOS boxes with differing affinities (Lewis *et al.* 1994). This allows variable basal expression of genes in the SOS response and timed induction of genes, after SOS is initiated (Janion 2001). Activation occurs by de-repression of the genes involved, which is achieved by the autolytic cleavage of LexA, in response to a signal mediated by the binding of RecA to ssDNA (Higashitani *et al.* 1995). When damage occurs, ssDNA accumulates in the cell, onto which RecA can load to form a nucleoprotein filament. DSBs can lead to SOS induction as ssDNA is produced during end processing (Rinken and Wackernagel 1992; Kogoma 1997). When bound to ssDNA, RecA has been referred to as RecA\* and is the signal for LexA autolytic cleavage to occur. SOS regulated genes, including LexA and RecA, can then be transcribed. Although it has been reported that expression of around a thousand genes can be affected during SOS induction (Courcelle *et al.* 2001; Khil and Camerini-Otero 2002; Quillardet *et al.* 2003), only around forty are defined as members of the SOS regulon, i.e. are directly regulated by LexA (Fernandez De Henestrosa *et al.* 2000; Courcelle *et al.* 2001). As DNA damage is repaired, the abundance of ssDNA and RecA\* decreases, therefore the newly synthesised LexA can bind once again to SOS boxes and repress the SOS response (Walker 1996). During prolonged damage, factors such as SulA are induced, which inhibits cell division until repair can be completed (Trusca *et al.* 1998).

### **1.5.2 The role of RecN as an SOS induced DSBR protein**

Induction of the SOS response drastically improves survival of cells exposed to UV and IR. For instance, when the SOS response is induced using a low dose of IR and time allowed for protein synthesis, cells were noted to be much more resistant to subsequent doses of IR (Pollard and Achey 1975;

Krasin and Hutchinson 1981). This suggests DSBR is enhanced after SOS induction and it is tempting to speculate that this may involve RecN, as it is one of the genes which is heavily induced as part of the SOS response. In fact, from 2D gel analysis, RecN is one of only four proteins that can be seen to be SOS induced, and after RecA, it is the most abundant SOS induced protein (Finch *et al.* 1985a). This was confirmed by studies using a *lac* reporter fused to the *recN* promoter, showing the gene to be induced at least 8-fold upon SOS induction (Picksley *et al.* 1984b). Furthermore, work by Rostas *et al.* (1987) identified two *lexA* binding motifs in the *recN* promoter region. This is unusual; only two other SOS response genes, *lexA* itself and *ydjM*, have more than one LexA binding motif (Little *et al.* 1981; Fernandez De Henestrosa *et al.* 2000). Sequence analysis of the *recN* promoter region and codon usage supports RecN having a low basal expression, which can be heavily upregulated (Rostas *et al.* 1987). Recently, Erill *et al.* (2007) analysed the SOS response of species across the gamma proteobacteria. This study suggested that the SOS response is conserved across species and all the SOS regulons have a core set of induced proteins. They showed that *recN* belonged to this core set of genes along with key components of the SOS response, such as *recA* and *lexA*. The authors suggest that these gene products are integral to the SOS response of all the species analysed. They also identified a putative, third *lexA* binding site upstream of *recN*, although a regulatory role, if any, has yet to be determined (Erill *et al.* 2003; Erill *et al.* 2007).

RecN would appear to be a major player in the SOS response and an important factor in the repair of DSBs, and this would explain the sensitivity of *recN* mutants to IR (Sargentini and Smith 1983). Exposure to IR causes DNA fragmentation, which is far more pronounced in a RecN deficient strain, than in wild-type bacteria. It has been suggested that RecN could limit fragmentation by aiding repair, or reducing the occurrence, of DSBs (Picksley *et al.* 1984a). When analysed further, RecN deficient strains were shown to be sensitive, not only to IR, but also to other DNA damaging agents particularly compounds which cause cross-linking, like mitomycin C (Dye and Ahmad 1995) and bleomycin, which causes DSBs (Povirk *et al.* 1977). A *recN* mutant is actually

as sensitive to bleomycin as *recA* or *recBCD* mutants, which are effectively incapable of HR. The high sensitivity to bleomycin may represent a variable reliance on RecN for DSB repair, depending on the nature of a break (Picksley *et al.* 1984a; Kosa *et al.* 2004). In *E. coli*, mutation of *recN* causes a moderate induction of the SOS response. Thus it would seem that in the absence of functional RecN, DNA damage accumulates, suggesting that RecN has roles in DNA repair, besides its role in the SOS response (Simic *et al.* 1991; Dunman *et al.* 2000). Strains of *E. coli* carrying *recN* mutations also show a slightly elevated sensitivity to UV and to the alkylating agent methyl methanesulphonate (Sargentini and Smith 1983; Picksley *et al.* 1984a). Both primarily cause single stranded DNA lesions, although, there is evidence that UV can cause low levels of DSBs (Bonura and Smith 1975) and it is possible that MMS has a similar effect.

Meddows *et al.* (2005) confirmed a role for RecN in the repair of DSBs by using the yeast homing endonuclease I-SceI to generate inducible, site-specific DSBs. This study showed that RecN had a moderate sensitivity to a single DSB, however recombination deficient strains carrying *recA* and *recBCD* mutations were far more sensitive. It was only when cleavage sites were introduced at several chromosomal loci that a *recN* strain became acutely sensitive to I-SceI (Meddows *et al.* 2005). Analysis of the repaired break sites also showed that, in the absence of RecN, deletions around the cleavage site were more common and more extensive. This led to the speculation that RecN not only has a role in the repair of DSBs, but also in the accuracy of their repair (Meddows *et al.* 2005). It has since been shown that the moderate phenotype of *recN* to a single DSB was at least in part due to the frequency of cleavage of the cleavage site used in this study. When a more active cleavage site was introduced, *recN* mutants become incredibly sensitive to I-SceI exposure, with survival comparable to *recA* and *recBCD* mutants (Grove *et al.* 2008).

### **1.5.3 RecN is involved in numerous recombination processes**

Both the RecBCD and RecFOR complexes can promote recombination, therefore RecN could function in either pathway. To determine in which RecN functions, the moderate UV sensitivity of a *recN* mutant was exploited. Genetic



studies combining *recN* mutations with *recB* and *recD* mutations, had no effect, but when a *recN* mutation was introduced into a *recF* background, a notable increase in UV sensitivity was observed. This was also the case, though the sensitivity was less marked, when *recN* mutations were combined with mutations in another RecFOR pathway gene, *recJ*. These data all suggested RecN was in the RecBCD recombination pathway (Lloyd *et al.* 1988; Wang and Smith 1988). Wang and Smith (1988) investigated this further, using neutral alkaline sucrose gradients to distinguish between DSBs and unrepaired daughter strand gaps, which are normally repaired by the RecFOR pathway. As expected for a RecBCD pathway component, RecN was important for the repair of UV induced DSBs, but had no effect on the repair of daughter strand gaps (Wang and Smith 1988). However, a *recN* strain that also carries both *recBCD sbcBC* mutations and should therefore be capable of recombination, is actually recombination defective, suggesting that in this background recombination requires a functional *recN* gene product (Picksley *et al.* 1984a).

Peterson and Mount (1993) showed that *dam recN*, and *dam recBC sbcBC* mutant strains (so only the RecFOR pathway is functional), are viable, but a *dam recN recBC sbcBC* mutant is inviable (Peterson and Mount 1993). Additionally, when *recD recJ* double mutants have a *recN* mutation introduced they show increased UV sensitivity and reduced levels of recombination (Lloyd *et al.* 1988). The conflicting information led to the suggestion that perhaps there are three epistatic recombination groups in *E. coli*; RecF, RecBCD and RecN (Lloyd and Buckman 1991). This notion of placing the *recN* gene in its own separate epistatic group is supported by work in *B. subtilis*, where *recN* forms one of the seven separate epistatic groups involved in recombination (Sanchez *et al.* 2007b).

RecN is also involved in recombination-based processes like conjugation. In a *recN* mutant DNA transfer is almost completely eliminated (Lloyd and Buckman 1995) and it was proposed that RecN was required to protect the incoming ssDNA from nucleases. This idea is supported by evidence from *B. subtilis*, which is naturally competent and takes up DNA

under certain growth conditions. Kidane and Graumann (2005) showed that during periods of competent growth, a RecN-YFP fusion protein formed foci, localised to the cell poles, in the same region as proteins associated with uptake of DNA. These foci dissipated after DNA uptake.

Like many proteins involved in recombination, RecN also appears to have a role in DNA replication. Experiments in *E. coli*, using a non-replicative plasmid with an endonuclease cleavage site showed that RecN was crucial for propagation of the plasmid. Induction of the endonuclease results in cleavage of a small fraction of the plasmid population, which generates a DSB. This can be processed and invasion of an intact plasmid generates a D-loop, allowing initiation of replication. The requirement for RecN could be eliminated if cells were SOS induced (Asai *et al.* 1994). When a *recN* mutation was combined with a *polA* mutation, cells showed increased sensitivity to UV and IR, suggesting that RecN, like many of the recombination proteins, perhaps has a role in stabilising or rebuilding collapsed replication forks (Sargentini and Smith 1983).

#### **1.5.4 The activity of RecN *in vitro***

RecN is widely distributed throughout the bacterial kingdom and the proteins show a high degree of functional conservation. In *Haemophilus influenzae* the RecN protein is a key component of this organism's SOS response (Sweetman *et al.* 2005). In *Deinococcus radiodurans* and *Helicobacter pylori*, *recN* mutants demonstrate phenotypes comparable to those of *E. coli recN* mutants, including sensitisation to IR, UV and mitomycin C exposure (Funayama *et al.* 1999; Wang and Maier 2008). *H. pylori recN* mutants also show reduced virulence (Wang and Maier 2008).

Kidane and Graumann (2005) were the first to publish details of an *in vitro* activity for a RecN protein, showing that a His-tagged *B. subtilis* RecN protein (BsRecN) could bind ssDNA. This supported suggestions that RecN may function to protect ssDNA during processes such as conjugation and competence (Lloyd and Buckman 1995; Kidane and Graumann 2005). *B. subtilis* RecN (BsRecN) shares a high degree of homology with its *E. coli* counterpart. It is LexA regulated and mutants are sensitive to agents like

mitomycin C and show reduced levels of recombination (Van Hoy and Hoch 1990; Alonso *et al.* 1993). The binding of BsRecN to ssDNA was reported to be ATP independent (Sanchez and Alonso 2005). However, in the presence of ssDNA (preferentially linear ssDNA), RecN showed increased ATPase activity, suggesting that ATP does affect DNA binding. If  $Mg^{2+}$  ions and either ADP or ATP were added to the reactions, BsRecN formed higher order structures with ssDNA, which were broken down by the addition of RecA. Their formation is also inhibited by the presence of SSB (Sanchez and Alonso 2005). This was supported by AFM binding studies that showed large accumulations of ssDNA-BsRecN protein complexes *in vitro*, which could also be disassembled by RecA (Sanchez *et al.* 2007a).

The year before the Kidane and Graumann (2005) study, wild-type BsRecN was partially purified and shown to run as an octamer on gel filtration, with smaller, dimeric and tetrameric species detectable by glycerol gradient centrifugation (Kidane *et al.* 2004). Higher order RecN complexes could even be detected *in vivo*, after DSBs had been induced. These structures were postulated to represent large nucleoprotein complexes that formed in response to DSBs (Kidane *et al.* 2004). The same study made use of fluorescently tagged *B. subtilis* RecN, RecF and RecO homologs and demonstrated that they were all recruited to the nucleoid, in an sequential manner, after DNA damage, forming one or two foci per cell. RecN was the first protein recruited to the nucleoid, and if absent the RecF and RecO foci did not form. It was therefore suggested that perhaps RecN could bind and ‘flag’ a DSB, aiding in the recruitment of subsequent repair proteins (Kidane *et al.* 2004). BsRecN also formed nucleoid-associated foci in the absence of SOS induction, suggesting the basally expressed RecN has a role perhaps aiding repair of breaks that arise due to replication fork collapse. Binding to them as it does other DSBs (Kidane *et al.* 2004). Consequently, perhaps RecN has a role similar to that of the eukaryotic MR(N/X) complexes (Kidane *et al.* 2004; Sanchez *et al.* 2007a). These complexes are believed to be early sensors of a DSB, recruiting other repair proteins and having active roles in the repair of a DSB (Usui *et al.* 1998; Mirzoeva and Petrini 2001; Usui *et al.* 2001; McGowan and Russell 2004).

Similarly, *E. coli* RecN was shown to be recruited to the nucleoid, after DNA damage (Nagashima *et al.* 2006; Moore Unpublished data). However, the tagged *E. coli* RecN also formed large cytoplasmic protein aggregates, which dissipated after the damage was repaired (Nagashima *et al.* 2006).

If RecN does act as a flag of DSBs it would be expected that it would be one of the first proteins recruited to a break site, prior to end-processing. In *B. subtilis* this appears to be true; a deletion of *recJ* and *addAB* (the *recBCD* homolog) prevents end processing. Yet, BsRecN-YFP foci still formed on the nucleoid in this strain and were actually more common (Sanchez *et al.* 2006), leading the authors to propose that RecN binds to a DSB, which cannot be repaired in this background and therefore RecN remains bound to the unprocessed breaks, thus explaining the multiple foci. However, contrary to this hypothesis is BsRecN's *in vitro* activity, where it has been shown to bind ssDNA, preferentially ssDNA with a 3'-OH group, which would only present after end processing of a break (Kooistra *et al.* 1993; Sharples and Lloyd 1993; Sanchez and Alonso 2005). The following questions must therefore be resolved: if BsRecN binds ssDNA, which is only present at a break site after end-processing rather than duplex DNA, how then does RecN bind and flag a break prior to end processing? Furthermore, what is the reason for the differences in the observed localisation of *E. coli* and *B. subtilis* RecN (Kidane *et al.* 2004; Nagashima *et al.* 2006).

The inherent insolubility of *E. coli* RecN has proven problematic for biochemical studies of this protein (Meddows 2002). However, *E. coli* RecN has been demonstrated to be a substrate of the ClpXP protease (Neher *et al.* 2006), which is composed of two subunits and degrades proteins in an ATP-dependent manner (Kim and Kim 2005). ClpXP has numerous targets including a large number of SOS induced proteins, such as UvrA, UmuD, LexA cleavage fragments (formed when LexA undergoes autolytic cleavage) and RecN (Neher *et al.* 2003; Flynn *et al.* 2004; Neher *et al.* 2006). RecN has a seven amino acid ClpXP tag at its C-terminus, which targets RecN to ClpXP. Due to the activity of ClpXP, RecN has a half-life of approximately 8 minutes. RecN is heavily induced, within 20 minutes of the SOS response being

initiated, but is completely degraded within 160 minutes, which corresponds to the appearance and disappearance of GFP-foci in the nucleoid and the large fluorescent protein aggregates seen in the cytoplasm (Nagashima *et al.* 2006; Neher *et al.* 2006). Mutation of the ClpXP-tag on RecN prevents targeting to ClpXP and drastically increases the half-life of RecN molecules. Consequently, the protein aggregates remain in the cytoplasm and cells show an increased sensitivity to DNA damage (Nagashima *et al.* 2006). A similar DNA damage sensitisation was observed when wild-type *E. coli* RecN was overexpressed from a plasmid (Meddows 2002). It is plausible that the accumulation of RecN, perhaps as cytoplasmic aggregates, is toxic to the cell.

RecN has been compared to the MR(N/X) complex of eukaryotes, functioning as a DSB sensor (Sanchez *et al.* 2007a). One of the key components of the MR(N/X) complex is Rad50, a member of the structural maintenance of chromosome (SMC) protein family (Hopfner *et al.* 2000). Sequence analysis of *E. coli* RecN, *Pyrococcus furiosus* Rad50 and SbcC, an *E. coli* SMC family member, revealed that they are all structurally related (Sharples and Leach 1995), suggesting that RecN is also a member of this protein family and could share their properties (Gorbalenya and Koonin 1990; Sharples and Leach 1995; Meddows 2002).

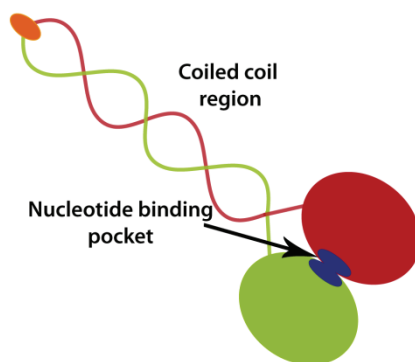
### **1.5.5 RecN is an SMC like protein**

SMC proteins are nucleoid-associated ATPases found in most organisms. They have diverse roles in DNA metabolism including DNA repair, chromosome cohesion and chromosome condensation (Hirano 2002). SMC proteins have a distinctive structure (Figure 1.9); a long coiled-coil region separates their two globular domains, one at each terminus of the protein. The coiled-coil region can fold back on itself, so as to bring the globular domains together. The folding is facilitated by a central flexible-hinge domain (Graumann 2001; Haering *et al.* 2002). The globular N-terminal and C-terminal domains possess Walker A and B box motifs, respectively. When brought together, these two motifs form a functional nucleotide binding and hydrolysis pocket (Walker *et al.* 1982; Hopfner *et al.* 2000).

A more in-depth sequence analysis was performed by Meddows (2002), who compared RecN once again to Rad50 of *Pyrococcus furiosus* and the SbcC of *E. coli*. Despite the much smaller size of RecN (*E. coli* RecN is 553 amino acids long, while *P. furiosus* Rad50 is 1313 amino acids and SbcC is 1048 amino acids), there was a high degree of homology, particularly in the head domains around the Walker A and B boxes. However, the author noted that a glutamate residue, postulated to be essential for the binding of  $Mg^{2+}$ , which is thought to be required for the hydrolysis of ATP (Hopfner *et al.* 2000), was not present in RecN.



**Figure 1.9. Generalised structure of an SMC protein.**

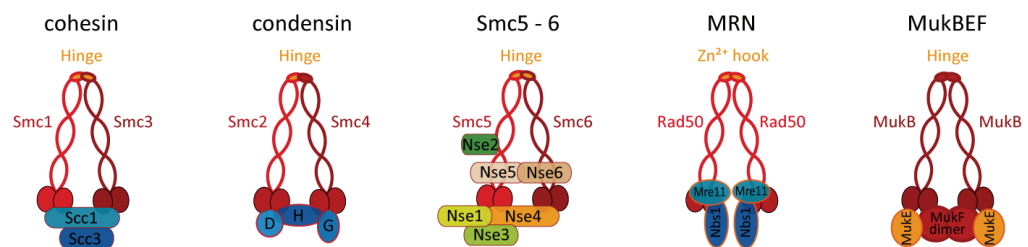


The two globular domains at the N- and C-terminus, (dark green and dark red respectively) are separated by a long coiled-coil region, these are brought together via folding of the molecule at the hinge domain (orange region). The Walker A and B boxes, which make up the nucleotide-binding pocket, are brought together by the folding of the molecule.

Eukaryotes possess multiple SMC proteins and at least six, named Smc1 – 6 respectively, are essential. They form heterodimers, through interactions at their hinge domains (Haering *et al.* 2002; Hirano and Hirano 2002), and are core components of three DNA-associated SMC-containing complexes, namely cohesin, condensin and the Smc5 – 6 complex (Figure 1.10) (Hirano 2002; Losada and Hirano 2005). These all have roles in chromosome dynamics and the repair of damaged DNA (Hirano 2002; Lehmann 2005; Strom and Sjogren 2007).

Cohesin is responsible for the tethering together of sister chromatids produced during DNA replication (Figure 1.10) (Nasmyth and Haering 2005). This tethering, referred to as sister chromatid cohesion, is essential to allow the correct segregation of chromosomes during mitosis (Tanaka *et al.* 2000). Mutations of the cohesin complex components leads to missegregation of

chromosomes and have also been linked to cancer and other human diseases (Sjogren and Nasmyth 2001; Tonkin *et al.* 2004; Wang *et al.* 2004). It is proposed that cohesin forms a ring-like structure (see Figure 1.10) with its two, non-SMC components linking the SMC protein heads. This structure could encircle DNA, trapping it inside the complex and potentially hold the sister chromatids together (Ciosk *et al.* 2000; Haering *et al.* 2002; Gruber *et al.* 2003; Ivanov and Nasmyth 2005; Ivanov and Nasmyth 2007). These rings are incredibly stable and loading is facilitated by opening of the ring at the SMC hinge, in response to ATP hydrolysis in the head domains (Hirano *et al.* 2001; Arumugam *et al.* 2003; Gruber *et al.* 2006; Hirano and Hirano 2006; Shintomi and Hirano 2007). This ATPase activity is stimulated by the non-SMC subunits (Arumugam *et al.* 2006). Although the entrapment model is not the only model to explain SMC-DNA interactions, it is the most widely accepted and the basis for our understanding of how SMC complexes interact with DNA (Hirano 2002; Huang *et al.* 2005; Guacci 2007). Loss of sister chromatid cohesion actually requires enzymatic cleavage of the complex, which allows separation of chromosomes at anaphase (Tomonaga *et al.* 2000; Uhlmann *et al.* 2000).



**Figure 1.10. Proposed model structures of the known SMC protein complexes.**

Cohesin, condensin and Smc5 – 6 are discussed in detail in the text. Each forms a heterodimer, with the heads bridged by a non-SMC protein. The Smc5 – 6 complex is unique in that the head is bridged by Nse4, while a sub-complex of Nse5/6 can bridge the heterodimer within the coiled-coil region. The MRN complex acts to flag DSBs, in some organism Xrs2 replaces Nbs1. MukBEF are the *E. coli* SMC structural homolog, showing functional homology to condensin.

Condensin (Figure 1.10) forms a complex much like cohesin. It is responsible for compacting chromosomes (Hirano 2002; Hirano 2005a), which is important for their segregation. The complex also has roles in transcription,

DNA repair, and in vertebrates it interacts with components of the BER pathway (Figure 1.1) (Chen *et al.* 2004; Blank *et al.* 2006; Heale *et al.* 2006). The third complex, Smc5 – 6, is primarily involved in DNA repair, but also the segregation of repetitive regions of DNA (Torres-Rosell *et al.* 2005). Initially identified in *S. pombe* as mutations that were IR and UV sensitive (Nasim and Smith 1975; Phipps *et al.* 1985), the Smc5 and Smc6 proteins were subsequently shown to be SMC proteins and part of a third, essential eukaryotic SMC complex (Fousteri and Lehmann 2000; Lehmann 2005; Sergeant *et al.* 2005). Interestingly, not all of the non-SMC components of the complex are essential, at least in *S. pombe*, although they are in *S. cerevisiae*. The Nse subunits of the complex also appear to interact in a novel manner, with several interacting with the SMC protein within the coiled-coil region (Figure 1.10) (Fousteri and Lehmann 2000; Sergeant *et al.* 2005; Palecek *et al.* 2006). These subunits are also unusual as they have distinct biochemical activities and are not just structural components of the complex (Andrews *et al.* 2005; Potts and Yu 2005; Zhao and Blobel 2005; Potts and Yu 2007).

Interestingly, both cohesin and the Smc5 – 6 complex have been implicated in DNA segregation, HR, replication and DNA repair, particularly of DSBs. Mutations in either complex cause mild UV sensitivity and a more pronounced sensitivity to IR (Nasim and Smith 1975; Phipps *et al.* 1985; Birkenbihl and Subramani 1992). This would suggest a defect with the processing of DSBs and there is a surfeit of evidence suggesting both complexes are involved in the repair of DSBs by HR (Mengiste *et al.* 1999; Skibbens *et al.* 1999; Sjogren and Nasmyth 2001; Morishita *et al.* 2002; Lehmann 2005; Strom and Sjogren 2007). Both complexes also show differential recruitment to damaged chromosomes and there is evidence that they can be recruited to break sites in an Mre11-dependent manner (Kim *et al.* 2002a; Strom *et al.* 2004; Unal *et al.* 2004; De Piccoli *et al.* 2006; Lindroos *et al.* 2006; Potts *et al.* 2006). Mre11 is part of the MR(N/X) DSB sensor complex and also contains another SMC protein, Rad50 (Figure 1.10). Exactly how these two complexes are involved in HR is unclear, but they seem to be very important. In their absence, chromosomal crossovers either do not occur



or occur abnormally (Klein *et al.* 1999; Mengiste *et al.* 1999; Buonomo *et al.* 2000; Lehmann 2005; Strom and Sjogren 2007). It has been speculated that the cohesin and Smc5 – 6 complex could aid HR by holding homologous sister chromatids in close proximity improving strand exchange (Sjogren and Nasmyth 2001).

These SMC complexes appear to have a lot in common with RecN, besides their apparent structural similarities (Sharples and Leach 1995; Meddows 2002). They are also involved in HR, particularly the repair of DSBs and probably the maintenance of replication forks. There are known bacterial SMC proteins, which share structural and functional homology with those discussed. *E. coli* MukBEF proteins form a complex with structural similarity to these SMC complexes (Figure 1.10) (Niki *et al.* 1992; Bartosik and Jagura-Burdzy 2005). Although, the role of MukBEF remains contentious as it has been suggested to facilitate both chromosome condensation (Ohsumi *et al.* 2001), and has also been proposed to function akin to cohesin (Sunako *et al.* 2001). However, it's essential role is apparent from the phenotype of a *mukB* mutant, in which chromosomes fail to partition correctly (Niki *et al.* 1992). Known SMC homologs, with apparent roles in chromosome segregation also exist in *Thermatoga maritima* and *B. subtilis* (Hiraga 2000; Lowe *et al.* 2001). RecN has also been compared to the SMC containing MR(N/X) complex (Sanchez *et al.* 2007a). Although, the Rad50 component shares homology with another *E. coli* protein, SbcC. Many functions have been assigned to Rad50 containing complexes (reviewed in Haber 1998a) and the SbcCD complex also appears to have numerous cellular roles.

It appears that *E. coli* cells possess proteins that can fulfil all the potential roles that RecN has been implicated in. However, unlike the other SMC homologs, RecN has one key distinguishing feature, being a member of the SOS regulon. This suggests that it is only required during extreme circumstances, perhaps functioning like cohesin that can induce genome wide sister chromatid cohesion in response to DNA damage (Kim *et al.* 2002b; Strom and Sjogren 2005). Despite this it does seem likely that RecN has some

role in the day to day maintenance of the cell, as a *recN* mutant is mildly SOS induced in the absence of exogenous DNA damage (Dunman *et al.* 2000).

## 1.6 Summary

Damage to DNA is probably the major threat to an organism's survival and is unavoidable, often caused by endogenously generated agents. A cell must therefore either tolerate or repair damaged DNA. A DSB is one of the most genotoxic lesions that can occur, it inhibits replication and due to its recombinogenic nature causes genomic instability. Repair of a DSB is therefore essential.

RecN is one of several bacterial proteins implicated in the repair of DSBs and more generally in HR. The role RecN fulfils may depend on the level of expression, requiring high levels of RecN to repair chronic damage to the genome, while basal expression is adequate to support HR processes, like conjugation, competence and perhaps the processing of collapsed replication forks. In the absence of RecN, lesions cannot be correctly processed and this leads to chronic SOS induction, which in turn leads to mutagenesis and genomic instability (Higashitani *et al.* 1995; Mukherjee *et al.* 1998). Therefore, by aiding repair and limiting the duration of the SOS response, RecN helps ensure genomic stability and cell survival (Savory 2007). Studies of RecN have suggested that it may have a dual role in the DSB repair pathways, acting before end processing, as a DSB sensor, and after end processing, when it could bind to the ssDNA present. However, neither role has been confirmed and biochemical data on RecN remains incomplete

RecN shows structural homology to SMC proteins (Sharples and Leach 1995), a diverse family of proteins with roles in DNA metabolism, including chromatid segregation, DNA repair, and HR (Hirano 1999). Loss of these SMC proteins is usually lethal and mutations that are viable cause gross chromosomal abnormalities and severe defects in chromosome segregation (Hirano 2005b). This contrasts with what we know about RecN; it is not essential and during normal growth, the absence of RecN has an insignificant effect on a cell. It is only after DNA damage occurs that the deleterious nature

of RecN deficiency becomes apparent. Cells lacking RecN are extremely sensitive to DNA lesions, such as DSBs, which require HR to facilitate repair. The major importance, yet non-essential role, of RecN to HR makes it rather unique. Perhaps RecN, which is heavily SOS induced, could provide cohesion of sister chromatids to facilitate repair, akin to what is seen with eukaryotic cohesin after DNA becomes damaged. Further study of RecN may improve our understanding of both HR and SMC proteins.

In this thesis, numerous approaches were taken to unravel the mystery that still surrounds RecN's function *in vivo*. This includes a continuation of the genetic studies, making use of a synthetic lethality assay to isolate any mutations that are lethal in conjunction with a *recN* mutation. Although no such mutants were isolated, this study did prove fruitful, identifying interactions between DSB repair pathways and the regulation of oxidative stress within a cell. The second approach taken was based on biochemical analysis. Although *E. coli* RecN proved non-amenable to purification, I will report details on the isolation and purification of RecN proteins from several bacterial species, including one homolog that can function in *E. coli*, fully complementing a deletion of the *recN* gene. These purified proteins have been shown to have *in vitro* activities, including ATPase and DNA binding properties, and their potential to provide structural data was assessed, including attempts to crystallise the proteins and identification of discrete protein complexes under an electron microscope. This provides new avenues for the study of RecN, and the ubiquitous SMC proteins.

# Chapter 2

## Materials and methods

### 2.1 Materials

#### 2.1.1 Chemicals and other Materials

The majority of the analytical reagent grade chemicals used in this study were purchased from Sigma, Fisher, Fluka and Melford. Agarose was purchased from Flowgen, acrylamide stock solution (acrylamide/bis-acrylamide 29:1) from Severn Biotech Ltd., and scintillation fluid from Packard. Exceptions are detailed in the relevant sections. The water used in this study was deionised and filtered with a USF ELGA Option 7/15 water purification unit. Where necessary, it was sterilised by autoclaving at 121°C for 15min.

#### 2.1.2 Radiochemicals

[ $\alpha^{32}\text{P}$ ] ATP was purchased from GE Healthcare at 5000Ci/mmol.

#### 2.1.3 Enzymes and other Proteins

All of the restriction endonucleases were purchased from New England Biolabs (NEB). As were T4 polynucleotide kinase, Calf Intestinal Alkaline Phosphatase (CIP) and T4 DNA ligase. Red Hot *Taq* DNA polymerase was obtained from Abgene, while Dynazyme EXT DNA polymerase and Phusion DNA polymerase were from NEB (FINNZYMES). Unless otherwise stated, all the enzymes were used with the buffers supplied in accordance with the supplier's recommendations. NTPs (Invitrogen) for PCR were diluted to a final concentration of 10mM (2.5mM of each NTP) using sterile distilled water and. ATP used in assays was made at 100mM stock concentration in dH<sub>2</sub>O and adjusted to pH7.0. Both were stored in small aliquots at -20°C

#### 2.1.4 Buffers and solutions

All the concentrations given are for final working concentration, stock concentrations, when made, are detailed.

### **DNA buffers**

- Ficoll loading buffer – 3% (w/v) Ficoll 400, 0.05% (w/v) bromphenol blue/xylene cyanol (stock solution, 5x).
- GBB (Gel Binding Buffer) – 50mM Tris-HCl pH8.0, 5mM EDTA, 1mM DTT, 100µg/ml BSA, 6% (v/v) glycerol (stock solution, 5x).
- HB (Helicase Buffer) – 20mM Tris-HCL pH7.5, 2mM DTT, 100µg/ml BSA.
- LIS – 6.7mM Tris-HCl pH8.0, 3.3mM sodium acetate, 2mM EDTA pH8.0 (stock solution, 10x).
- SSC – 150mM NaCl, 15mM sodium citrate, pH7.0 (stock solution, 20x).
- SSPE – 150mM NaCl, 3.6mM sodium dihydrogen orthophosphate monohydrate, 1mM EDTA (stock solution, 20x).
- TAE – 40mM Tris-acetate, 1mM EDTA (stock solution, 50x).
- TE – 10mM Tris-HCl pH7.5, 1mM EDTA.
- TBE – 90mM Tris-borate, 2mM EDTA (stock solution, 10x).
- TNE – 10mM Tris-HCl pH8.0, 10mM NaCl, 10mM EDTA.

### **Protein buffers**

- Buffer A (standard low salt buffer) – 50mM Tris-HCl pH7.5, 1mM EDTA, 1mM DTT.
- Buffer B (standard high salt buffer) – As above plus 1M NaCl.
- Gel Filtration buffer – 50mM Tris-HCl pH7.5, 150mM NaCl.
- Lysis buffer – 50mM Tris-HCl, pH7.5, 150mM NaCl.
- SDS PAGE running buffer – 0.1% (w/v) SDS, 1.44% (w/v) glycine, 0.3% (w/v) Trizma base (stock solution, 10x).
- SDS PAGE loading buffer – 50mM Tris-HCl pH6.8, 100mM DTT, 2% (w/v) SDS, 0.1% (w/v) bromphenol blue, 10% (v/v) glycerol (stock solution, 5x).
- Western transfer buffer – 25mM Tris-HCl pH8, 192mM glycine, 20% (v/v) methanol.

### **Miscellaneous Buffers**

- MC buffer – 100mM MgSO<sub>4</sub>, 5mM CaCl<sub>2</sub>

- Lambda buffer – 6mM Tris-HCl pH7.2, 10mM MgSO<sub>4</sub>, 0.005% (w/v) gelatine.
- TFB-1 - 100mM KCl, 50mM MnCl<sub>2</sub>, 30mM KAc, 10mM CaCl<sub>2</sub>, 15% glycerol (v/v).
- TFB-2 – 10mM Mops, 10mM KCl, 75mM CaCl<sub>2</sub>, 15% glycerol (v/v).
- CTAB Extraction solution – 100mM Tris-HCl pH7.5, 25mM EDTA, 2M NaCl.
- CTAB dilution solution – 50mM Tris-HCl pH7.5, 10mM EDTA, 1% Cetyl triethylammonium bromide (CTAB).

### 2.1.5 Growth media

Yeast extract, tryptone, bactoagar and MacConkey agar base were all obtained from Difco. Liquid and solid media used for the growth of *E. coli* strains were prepared according to the standard recipes below. Growth media were sterilised by autoclaving at 121°C for 15min.

- Luria and Burrows (LB) Broth and Agar – 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.05% (w/v) NaCl and 0.008% (w/v) NaOH, pH7.5. LB broth was supplemented with 1.5% (w/v) bactoagar to make LB agar.
- Luria-Bertani Broth – 1% (w/v) typtone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl and 0.008% (w/v) NaOH, pH7.5.
- Mu Broth and Agar – 1% (w/v) typtone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl and 0.008% NaOH, pH7.5. Mu broth was supplemented with 1% (w/v) bactoagar to make Mu agar. Overlays were made by supplementing Mu broth with bactoagar to either 0.4% (w/v) or 0.6% (w/v).
- Minimal 56/2 Salts – 78mM KH<sub>2</sub>PO<sub>4</sub>, 98mM NaHPO, 0.01% MgSO<sub>4</sub>·7H<sub>2</sub>O (w/v), 0.1% (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (w/v), 0.0005% Ca(NO<sub>3</sub>)<sub>2</sub> (w/v) and 0.000025% FeSO<sub>4</sub>·7H<sub>2</sub>O (w/v).
- Minimal salts agar medium – made using double strength 56/2 salts media supplemented with glucose to 3.3mg/ml and thiamine to 1µg/ml for basic salts agar media. Nucleotides were added at 5-10µg/ml except for thymine, which was at 100µg/ml. Amino acids were added to a concentration of 50–80µg/ml as required for growth of auxotrophic strains. This was mixed with an equal volume of water with 3% (w/v) bactoagar to give 1.5% agar plates.

- P1 Agar – LB agar supplemented with 5mM CaCl<sub>2</sub> and 0.13% (w/v) glucose (3 ml 0.5 M CaCl<sub>2</sub>/l, 2ml 20% glucose/l)
- SOB broth – 2% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.05% (w/v) NaCl, 10ml 1M MgCl<sub>2</sub> and 10ml 1M MgSO<sub>4</sub>/litre of media.
- SOC broth – 2% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.05% (w/v) NaCl and 20mM glucose.
- Tryptone (TB) – 1% (w/v) tryptone, 0.5% (w/v) NaCl, the media was supplemented with 0.7% (w/v) bactoagar to produce TB overlay agar and 1% bactoagar to make TB agar.
- YT media – 1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl

### **2.1.6 Antibiotics**

Antibiotic stock solution were made in sterile distilled water and stored at 4°C, except for tetracycline and apramycin, which were stored at -20°C. All the antibiotics were obtained from Sigma, except for chloramphenicol, which was sourced from Melford. Occasionally, when large quantities of ampicillin were needed, as during fermenter runs, the antibiotic was dissolved in ethanol to minimise handling volume. Ampicillin (amp), apramycin (apra), carbenicillin (cb), kanamycin (kan) and trimethoprim (dhfr) were used at final concentrations of 40µg/ml, while chloramphenicol (cat) and tetracycline (tc) were used at final concentrations of 10µg/ml

### **2.1.7 Bacterial strains, bacteriophages and plasmids**

The allele designations follow standard nomenclature, except where no standards apply when conventions established in the Lloyd laboratory and elsewhere are applied (Table 2.1). Bacterial strains used in this study are listed in Table 2.2 (commercial and public domain) and Table 2.3 (laboratory strains). All genotypes are derived from MG1655 unless otherwise stated.

**Table 2.1. Description of alleles used in this study**

Allele	description
$\Delta lacZYA::FRT$	The <i>lac</i> operon genes are deleted and replaced with kanamycin gene flanked by FRT repeats. These repeats allow removal of the kan gene but the repeats themselves remain in the genome.
$\Delta attB::P_{BAD}::I-SceI$	I- <i>SceI</i> expression cassette I- <i>SceI</i> expression is regulated by the arabinose inducible promoter $P_{BAD}$ and this promoter and I- <i>SceI</i> gene construct is inserted into the chromosome at the <i>attB</i> loci (Meddows <i>et al.</i> , 2004).
$\Delta lacZ::I-SceI_{CCS4}$	I- <i>SceI</i> cleavage site inserted into the chromosome at the <i>lacZ</i> loci. The site is linked to a chloramphenicol resistance gene (Grove <i>et al.</i> 2008).
$\Delta recN::Hi-recN$ -kan	<i>E. coli</i> chromosomal construct where the <i>E. coli recN</i> genes coding sequence are deleted and replaced with the coding sequence of <i>H. influenzae recN</i> strain Rd KW20 (genomic DNA was kindly provided by Dr. G. Thomas, University of York). The <i>E. coli recN</i> promoter is left in place ensuring the gene is regulated, as is the native gene. Immediately 3' of the new coding sequence is a kan resistance gene as a marker. (For construction see Chapter 4).
$\Delta recN::Aq-recN$ -kan	As above except the <i>E. coli recN</i> is replaced by the coding sequence of the <i>recN</i> gene of <i>A. aeolicus</i> (Chapter 4). <i>A. aeolicus recN</i> was cloned from pTRM129 (Meddows, 2002)
$\Delta recN::Bf-recN$ -kan	As above except the <i>E. coli recN</i> is replaced by the coding sequence of the <i>recN</i> gene of <i>B. fragilis recN</i> (Chapter 4). <i>Bacteriodes fragilis</i> strain VPI2553 genomic DNA was acquired from the ATCC
$\Delta recN::recN_{K35A}$ -kan	As the above <i>recN</i> gene deletions, except the coding sequence of <i>E. coli recN</i> is replaced by the coding sequence of <i>E. coli recN</i> encoding a mutated version of the protein with the lysine residue at position 35 mutated to an alanine (Chapter 6).



$\Delta recN::Hi-recN_{K35A}$ -kan	As above, but in this case it is a mutated version of the <i>H. influenzae recN</i> encoding the K35A substitution inserted into the chromosome to replace the <i>E. coli recN</i> coding sequence (Chapter 6).
<i>ahpC</i> ::kan	This study. Mini Tn10 insertion carrying a kanamycin resistance gene inserted 98 base pairs from the start of the <i>ahpC</i> genes coding sequence. Generated during mutagenesis of strain AM1581 using the EZ::TN <KAN-2> Tnp Transposome kit <sup>TM</sup> (Chapter 2.2.7 and Chapter 3, strain SW1005)
$\Delta oxyR$ ::apra	This study. A full deletion of the <i>oxyR</i> gene coding sequence with an apramycin resistance cassette inserted in its place, generated using the method described by Datsenko and Wanner (2000). Using primers Oapr-1 and Oapr-2 (Table 2.6) to clone by PCR an apramycin resistance gene flanked by homology to the <i>oxyR</i> gene (SW1039).

**Table 2.2. Commercial and public *E. coli* strains used in this study.**

Strain	Relevant genotype	Source or derivation
MG1655	F <sup>-</sup> $\lambda^-$ <i>rph-1</i>	Bachmann (1996)
DH5 $\alpha$	F' <i>endA1 hsdR17</i> (r <sub>K</sub> -m <sub>K</sub> ) <i>supE44 thi-1 recA1 gyrA</i> (Nal) <i>relA1 <math>\Delta lacZYA</math>-argF U169 deoR</i> ( $\Phi$ 80 <i>lac</i> ( $\Delta lacZ$ ) M15)	Raleigh <i>et al</i> (1989)
BL21 (DE3)	F <sup>-</sup> <i>ompT hsdS dcm lon gal</i> (r <sup>-</sup> m <sup>-</sup> <i>E. coli</i> B strain) with DE3 a $\lambda$ prophage carrying the T7 RNA polymerase.	Studier and Moffat (1986)
W3110	IN ( <i>rrnD</i> – <i>rrnE</i> )1	Bachmann (1996)
TB <sub>28</sub>	$\Delta lacZYA$ ::FRT	Bernhardt and de Boer (2003)

**Table 2.3. Laboratory strains of *E. coli* used in this study**

Strain	Relevant genotype	Source or derivation
AM1581	$\Delta lacZYA$ ::FRT <i>recB</i> <sub>268</sub> ::Tn10, carrying pAM375 ( <i>recB</i> <sup>+</sup> <i>lacZYA</i> <sup>+</sup> )	Laboratory strain
DIM435	$\Delta lacZYA$ ::FRT $\Delta recN$ <sub>266</sub> <i>tyrA</i> ::Tn10	Laboratory strain
JIG509	$\Delta attB$ ::P <sub>BAD</sub> ::I- <i>SceI</i> $\Delta lacZ$ ::I- <i>SceI</i> <sub>CCS4</sub>	Grove <i>et al</i> (2008)
JIG619	$\Delta recN$ ::kan	Laboratory strain

JIG623	BL21(DE3) $\Delta recN::kan$ , carrying pLysS	Grove <i>et al</i> (2008)
JIG625	$\Delta recN::dhfr$	Grove <i>et al</i> (2008)
JIG628	$\Delta attB::P_{BAD}::I-SceI \Delta lacZ::I-SceI_{CCS4}$ $\Delta recN::dhfr$	Laboratory strain
JIG645	$\Delta recN::Hi-recN-kan$	Laboratory strain
JIG646	$\Delta attB::P_{BAD}::I-SceI \Delta lacZ::I-SceI_{CCS4}$ $\Delta recN::Hi-recN-kan$	Laboratory strain
JIG652	$\Delta recN::Aq-recN-kan$	Laboratory strain
JIG653	$\Delta attB::P_{BAD}::I-SceI \Delta lacZ::I-SceI_{CCS4}$ $\Delta recN::Aq-recN-kan$	Laboratory strain
JIG655	$\Delta recN::Bf-recN-kan$	Laboratory strain
JIG656	$\Delta attB::P_{BAD}::I-SceI \Delta lacZ::I-SceI_{CCS4}$ $\Delta recN::Bf-recN-kan$	Laboratory strain
JIG685	$\Delta recN::Hi-recN_{K35A}-kan$	Laboratory strain
JIG723	$\Delta attB::P_{BAD}::I-SceI \Delta lacZ::I-SceI_{CCS4}$ $\Delta recN::Hi-recN_{K35A}-kan$	Laboratory strain
JIG725	$\Delta recN::recN_{K35A}-kan$	Laboratory strain
JIG726	$\Delta attB::P_{BAD}::I-SceI \Delta lacZ::I-SceI_{CCS4}$ $\Delta recN::recN_{K35A}-kan$	Laboratory strain
JJ1017	$\Delta lacZYA::FRT \Delta recG::cat$ , carrying pJJ100 ( $recG^+$ , $lacZYA^+$ )	Laboratory strain
N1710	$F^- supF supE hsdR galK trpR metB1 lacY tonA$	Laboratory strain
N4279	$recA_{269}::Tn10$	Laboratory strain
N5747	$\Delta lacZYA::FRT \Delta ruvC::cat$ , carrying pAM372 ( $ruvC^+$ $lacZYA^+$ )	Laboratory strain
N5988	$\Delta lacZYA::FRT recB_{270}::kan$ , carrying pAM375 ( $recB^+$ $lacZYA^+$ )	Laboratory strain
N6121	$\Delta lacZYA::FRT recA_{269}::Tn10$ , carrying pAM383 ( $recA^+$ $lacZYA^+$ )	Laboratory strain
N6909	$\Delta rep::dhfr$	Laboratory strain
SW1005	$\Delta lacZYA::FRT recB_{268}::Tn10 ahpC::kan$ , carrying pAM375 ( $recB^+$ $lacZYA^+$ )	$ahpC$ mutant isolated (Chapter 3)

SW1018	<i>ahpC::kan</i>	P1.SW1005 × MG1655 to kan
SW1019	$\Delta lacZYA::FRT \Delta recN_{266} tyrA::Tn10$ <i>ahpC::kan</i> , carrying pSW101 ( <i>recN</i> <sup>+</sup> <i>lacZYA</i> <sup>+</sup> )	P1.SW1018 × SW1054 to kan
SW1020	$\Delta lacZYA::FRT ahpC::kan$	P1.SW1018 × TB <sub>28</sub> to kan
SW1030	$\Delta lacZYA::FRT recB_{268}::Tn10 ahpC::kan$ , carrying pAM375 ( <i>recB</i> <sup>+</sup> <i>lacZYA</i> <sup>+</sup> )	P1.SW1018 × AM1581 to kan
SW1039	$\Delta oxyR::apra$	Deletion of the <i>oxyR</i> gene (Table 2.1)
SW1040	$\Delta lacZYA::FRT \Delta oxyR::apra$	P1.SW1040 × TB <sub>28</sub> to apra
SW1041	$\Delta lacZYA::FRT \Delta recN_{266} tyrA::Tn10$ $\Delta oxyR::apra$ , carrying pSW101 ( <i>recN</i> <sup>+</sup> , <i>lacZYA</i> <sup>+</sup> )	P1.SW1040 × SW1054 to apra
SW1042	$\Delta lacZYA::FRT \Delta recN_{266} tyrA::Tn10$ $\Delta oxyR::apra ahpC::kan$ , carrying pSW101 ( <i>recN</i> <sup>+</sup> <i>lacZYA</i> <sup>+</sup> )	P1.SW1040 × SW1019 to apra
SW1054	$\Delta lacZYA::FRT \Delta recN_{266} tyrA::Tn10$ , carrying pSW101 ( <i>recN</i> <sup>+</sup> <i>lacZYA</i> <sup>+</sup> )	Transformation of DIM435 with pSW101.
SW1097	$\Delta lacZYA::FRT recB_{270}::kan \Delta oxyR::apra$ , carrying pAM375 ( <i>recB</i> <sup>+</sup> , <i>lacZYA</i> <sup>+</sup> )	P1.SW1040 × N5988 to apra
SW1100	$\Delta lacZYA::FRT recA_{269}::Tn10 \Delta oxyR::apra$ , carrying pAM383 ( <i>recA</i> <sup>+</sup> , <i>lacZYA</i> <sup>+</sup> )	P1.SW1040 × N6121 to apra
SW1101	$\Delta lacZYA::FRT recA_{269}::Tn10 ahpC::kan$ , carrying pAM383 ( <i>recA</i> <sup>+</sup> <i>lacZYA</i> <sup>+</sup> )	P1.SW1018 × N6121 to kan
SW1120	$\Delta lacZYA::FRT \Delta recG::cat \Delta oxyR::apra$ , carrying pJJ100 ( <i>recG</i> <sup>+</sup> <i>lacZYA</i> <sup>+</sup> )	P1.SW1040 × JJ1017 to apra
SW1124	$\Delta lacZYA::FRT \Delta recG::cat ahpC::kan$ , carrying pJJ100 ( <i>recG</i> <sup>+</sup> <i>lacZYA</i> <sup>+</sup> )	P1.SW1018 × JJ1017 to kan
SW1135	$\Delta lacZYA::FRT \Delta ruvC::cat \Delta oxyR::apra$ , carrying pAM372 ( <i>ruvC</i> <sup>+</sup> <i>lacZYA</i> <sup>+</sup> )	P1.SW1040 × N5747 to apra
SW1145	$\Delta lacZYA::FRT \Delta ruvC::cat ahpC::kan$ , carrying pAM372 ( <i>ruvC</i> <sup>+</sup> <i>lacZYA</i> <sup>+</sup> )	(P1.SW1018 × N5747 to kan)
TRM160	$\Delta recN_{266} tyrA::Tn10$	Lab Strain

**Table 2.4. Plasmid vectors used in this study**

Plasmid	Description	Source
pET3a	Protein over-expression vector, uses phage T7 promoter system, encodes kan resistance	Novagen
pET14b	Protein over-expression vector, uses phage T7 promoter system, encodes amp resistance, can be used to add a his-tag at N-terminus of protein.	Novagen
pET22a	Protein over-expression vector, uses phage T7 promoter system, encodes amp resistance	Novagen
pET28b	Protein over-expression vector, uses phage T7 promoter system, encodes kan resistance	Novagen
pGEM-7Zf(-)	High copy cloning vector (300-400 copies/cell) encodes amp resistance	Promega
pKD46	Low copy plasmid containing 2154nt of phage $\lambda$ Red ( $\gamma$ , $\beta$ and <i>exo</i> genes) expressed from $P_{BAD}$ , and having a temperature sensitive replicon (oriR101) and encodes amp resistance	Datsenko and Wanner (2000)
pLau17	$P_{BAD}$ - <i>gfp</i> . Can be used to make GFP fusions. Also the source of the arabinose inducible promoter $P_{BAD}$ .	Lau <i>et al</i> (2003)
pLysS	Helper plasmid used in protein overexpression. Encodes T7 lysozyme which acts to suppress T7 polymerase expression. Contains the cat gene encoding cm resistance	Novagen
pRC7	Unstable, low copy, <i>lacZYA</i> <sup>+</sup> , encodes amp resistance (for details see Figure 3.1A)	Bernhardt and de Boer (2004)
pT7-7	Protein over-expression vector, uses phage T7 promoter system, encodes amp resistance.	Tabor and Richardson (1985)
pUC19	Temperature sensitive high copy number (75 copies at 37°C), encodes amp resistance.	NEB

**Table 2.5. Plasmid constructs used in this study**

Plasmid	Description	Source
pAM372	The <i>ruvC</i> gene of <i>E. coli</i> cloned into pRC7 as an <i>EcoRI</i> – <i>HindIII</i> fragment. Expression of the <i>ruvC</i> gene is controlled by the plasmids <i>lac</i> promoter	Laboratory plasmid
pAM375	The <i>recB</i> gene of <i>E. coli</i> cloned into pRC7 as an <i>EcoRI</i> – <i>HindIII</i> fragment. <i>recB</i> expression is regulated by the <i>lac</i> promoter (for details see Figure 3.1 B).	Laboratory plasmid
pAM383	The <i>recA</i> gene of <i>E. coli</i> cloned into pRC7 as an <i>EcoRI</i> – <i>HindIII</i> fragment. Expression of the <i>recA</i> gene is controlled by the plasmids <i>lac</i> promoter.	Laboratory plasmid
pAPS108	pGEM-7Zf(-) based plasmid encoding cat resistance used during recombineering as PCR template	Savory, 2007
pJG16	As pAPS108 but with the cat resistance gene, replaced by kan resistance gene	Laboratory plasmid
pJG71	The <i>H. influenzae recN</i> gene cloned into pLau17 as an <i>EcoRI</i> – <i>HindIII</i> fragment. Allows arabinose inducible expression of HiRecN.	Laboratory plasmid
pJG74	pET22 with the 489 base pairs of the <i>E. coli recN</i> genes promoter amplified from genomic DNA as an <i>XbaI</i> – <i>NdeI</i> fragment and cloned into the complimentary sites.	Laboratory plasmid
pJG75	As pJG74 but with a kan resistance gene cloned into the <i>HindIII</i> restriction site immediately downstream of the cloned promoter. Used in the creation of heterologous strains (Chapter 4)	Laboratory plasmid
pJG83	As pJG75 with the <i>H. influenzae recN</i> gene PCR amplified from genomic DNA, cloned between the <i>E. coli recN</i> promoter and the kan resistance gene as a <i>NdeI</i> – <i>HindIII</i> fragment.	Laboratory plasmid
pJG94	The <i>B. fragilis recN</i> gene PCR amplified from genomic DNA and cloned into pT7-7 as an <i>NdeI</i> – <i>BamHI</i> fragment.	Laboratory plasmid
pJG95	As pJG75 with the <i>A. aeolicus recN</i> gene sub-cloned from pTRM129 as a <i>NdeI</i> – <i>HindIII</i> inserted between the <i>E. coli</i> promoter and the kanamycin resistance gene	Laboratory plasmid
pJG98	As pJG75 with the <i>H. influenzae recN</i> gene PCR amplified from genomic DNA, cloned between the <i>E. coli recN</i> promoter and the kan resistance gene as a <i>NdeI</i> – <i>HindIII</i> fragment.	Laboratory plasmid

pJG99	The <i>B. fragilis recN</i> gene sub-cloned from pJG94 as <i>NdeI</i> – <i>HindIII</i> fragment into the compatible sites of pLau17. Allows arabinose inducible expression of BfRecN.	Laboratory plasmid
pJG121	pET22 with the <i>A. aeolicus recN</i> gene amplified from pTRM129 cloned in as a <i>NdeI</i> – <i>HindIII</i> fragment, used to give a C-terminal His tag. Used for overexpression.	Laboratory plasmid
pJG122	Sub-clone from pSW116 of <i>A. aeolicus recN</i> with the K35A substitution as a <i>NdeI</i> – <i>HindIII</i> fragment cloned into the complimentary sites of pET22 to generate a C-terminal His-tag. Used for overexpression.	Laboratory plasmid
pJG123	As pJG121 with the <i>H. influenzae recN</i> gene cloned in as an <i>NdeI</i> – <i>HindIII</i> sites to give a C-terminal his tag. Used for overexpression.	Laboratory plasmid
pJJ100	The <i>recG</i> gene and its promoter were amplified from <i>E. coli</i> genomic DNA and cloned into pRC7 as a blunt ended product into the <i>ApaI</i> site.	Laboratory plasmid
pSW101	The <i>recN</i> gene and its promoter were amplified by PCR from <i>E. coli</i> MG1655 using primers RecN-F and RecN-2. It was cloned as a blunt ended ligation into the <i>ApaI</i> site of pRC7.	This study
pSW103	The C-terminal 489nt of <i>E. coli recN</i> amplified from genomic DNA using primers SWN3 and SWN4 to introduce <i>NcoI</i> and <i>BamHI</i> sites cloned into the complimentary sites of pET28. Used for overexpression.	This study
pSW104	The N-terminal 456nt of <i>E. coli recN</i> amplified from genomic DNA using primers SWN1 and SWN2 that introduce a <i>NdeI</i> and <i>BamHI</i> sites the fragment is cloned into the complimentary sites of pET22. Used for overexpression.	This study
pSW114	Quickchange™ mutagenesis of pJG71 using primers SWN42 and SNW43 to introduce K35A mutation into <i>H. influenzae recN</i> .	This study
pSW115	Quickchange™ mutagenesis of pTRM129 using primers SWN40 and SNW41 to introduce K35A mutation into <i>A. aeolicus recN</i> .	This study
pSW116	<i>A. aeolicus recN</i> sub-cloned from pTRM129 as a <i>NdeI</i> – <i>HindIII</i> fragment inserted into the complimentary sites of pJG74	This study
pSW117	Quickchange™ mutagenesis of pSW116 using primers SWN40 and SNW41 to introduce K35A mutation into <i>A. aeolicus recN</i> .	This study
pSW118	Quickchange™ mutagenesis of pJG83 using primers SWN42 and SNW43 to introduce K35A mutation into <i>H. influenzae recN</i> .	This study

pSW119	Quickchange <sup>TM</sup> mutagenesis of pTRM106 using primers SWN44 and SNW45 to introduce K35A mutation into <i>E. coli recN</i> .	This study
pSW121	The <i>E. coli recN</i> gene sub-cloned from pTRM106 as an <i>NdeI</i> – <i>Bam</i> HI fragment and cloned into the complimentary sites of pJG75.	This study
pSW122	Quickchange <sup>TM</sup> mutagenesis of pSW121 using primers SWN44 and SNW45 to introduce K35A mutation into <i>E. coli recN</i> .	This study
pSW123	Quickchange <sup>TM</sup> mutagenesis of pJG123 using primers SWN42 and SNW43 to introduce K35A mutation into <i>H. influenzae recN</i>	This study
pTRM106	pT7-7 with the <i>E. coli recN</i> gene cloned in as a <i>NdeI</i> - <i>Bam</i> HI fragment. Used for overexpression.	Meddows, 2002
pTRM129	pT7-7 with the <i>A. aeolicus recN</i> gene cloned in as a <i>NdeI</i> - <i>Bam</i> HI fragment. Used for over-expression.	Meddows, 2002

## 2.1.8 Oligonucleotides

Summary of the oligonucleotides used in this study and an outline of their usage. All were ordered from MWG and were used directly without further purification.

**Table 2.6. List of oligonucleotides used in this study.**

Primer Name	Sequence 5' to 3'	Notes
ARB-1	GGC CAC GCG TCG ACT AGT CAN NNN NNN NNN GC	Arbitrary primer used in first round PCR to sequence the location of chromosomal kanamycin insertions (See Figure 3.2).
ARB-2	GGC CAC GCG TCG ACT AGT CA	Binds at 5' of the PCR product generated by ARB-1, used in the secondary PCR to sequence the location of chromosomal kanamycin insertions (See Figure 3.2).
ELB37	ACG CTG CCG AAT TCT GGC TTG CTA GGA CAT GCT GTC TAG AGA CTA TCG AT	Anneals with RGL13, used to make duplex for use in gel retardation assays (Chapter 6).
ELB38	ACG CTG CCG AAT TCT GGC TTG CTA G	Anneals to RGL16 to produce a 5' overhang, for use in gel retardation assays (Chapter 6).
EzKan-1	AAG CTC TCA TCA ACC GTG GCG G	Primer binds at 3' of kanamycin resistance gene (See Figure 3.2).
EzKan-2	TTG GTT GTA ACA CTG GCA GAG C	Primer binds at 3' of kanamycin resistance gene, downstream of EzKan-1 (See Figure 3.2).
Oapr-1	CGT GAT CTT GAG TAC CTG GTG GCA TTG GCT GAA CAC CGC CAT TTT CGG CGT CAT GTG CAG CTC CAT CAG C	Anneals to 5' of apramycin resistance gene and adds 40 base pairs of sequence homologous to the first 40 base 5' of the <i>oxyR</i> genes coding sequence. Used in conjunction with Oapr-2 to generate a deletion of the <i>oxyR</i> gene and its replacement with an apramycin resistance cassette



Oap <sup>r</sup> -2	AGC GGT GAG CCA GGA CGA TAA ACC AGG CCA ATA GTG CGG CGT GGT TCC GGC CCG CCC AGA TAC AGA AAA GC	Anneals at the 3' of the apramycin resistance cassette and introduces 40 base pairs of sequence homologous to the 3' end of the <i>oxyR</i> gene.
RGL16	ATC GAT AGT CTC TAG ACA GCA TGT CCT AGC AAG CCA GAA TTC GGC AGC GT	Radiolabelled for use as ssDNA substrate in gel retardation assays (Chapter 6)
PM2	GAC ATG CT G TCT AGA GAC TAT CGA T	Anneals to RGL16 to produce a 3' overhang, for use in gel retardation assays (Chapter 6).
RecN-1	ATG TTG GCA CAA CTG ACC ATC AGC A	Anneals at the ATG start codon of <i>E. coli recN</i> , to give a flush DNA end. This study.
RecN-2	CG GGC CCG AGC AGT ACG	Anneals flush at the stop codon of <i>E. coli recN</i> . This study.
RecN-F	GAT GTC CCG GGT CGT CGC	Anneals upstream 55 base pairs from the ATG start codon of <i>E. coli recN</i> giving a flush end.
SWN01	GTT ACT CAT ATG TTG GCA CAA CTG ACC ATC AG	Anneals 5' of <i>recN</i> and introduces an <i>NdeI</i> site before the start codon.
SWN02	TAC GGA TCC TTA CTG CAG TAG AGA GGT TTC ATT GG	Anneals 3' of the <i>E. coli recN</i> ATG start codon and allows cloning of the first 456nt of <i>recN</i> introducing an <i>NdeI</i> site at the start codon.
SWN03	GAT CAT CCA TGG CAG AAG AAC TTG CAG AGC TGA TC	Anneals at the 5' of the <i>E. coli recN</i> 489nt from the terminus of the gene and introduces a <i>NcoI</i> site 5' of the gene sequence.
SWN04	GAT GGA TCC TTA CGC TGC AAG CAG TTC TTT CG	Anneals at the terminus of <i>E. coli recN</i> coding sequence and introduces a <i>HindIII</i> site after the gene sequence.
SWN05	CGA GGG ATC CTG TTG GCA CAA CTG ACC ATC AG	Anneals as SWN1 but introduces a <i>BamHI</i> site.
SWN06	GCA CTC GAG TTA CTG CAG TAG AGG TTT CAT TTG G	Anneals as SWN2 but introduces a <i>XhoI</i> site.
SWN07	GGT GGA TCC TGG CAG AAG AAC TTG CAG AGC TGA TC	Anneals as SWN3, but introduces a <i>BamHI</i> site
SWN08	GGT CTC GAG TTA GGC TGC AAG GCA GTT CTT TCG	Anneals as SWN4, but introduces a <i>XhoI</i> site.

SWN09	ATT CAT ATG GGT TAA AAG AGG TGC CGC TCC G	Anneals 3' of the <i>E. coli oxyR</i> genes promoter and introduces an <i>NdeI</i> site before the gene sequences.
SWN10	ATT GGA TCC CTA TCG GGT AGC TGC GTT AAA CGG	Anneals at terminus of the <i>E. coli oxyR</i> gene and introduces a <i>BamHI</i> site.
SWN15	CTA GGA TCC TTA TGT ATT AGG TGT GAC TTC ACT GCC ACC	Anneals 30 nucleotides before the end of the <i>E. coli recN</i> gene sequence, allows amplification of the gene without the last 10 amino acids which comprise the ClpXP tag.
SWN33	GGA AGG ACG TCA ATA GTC AC	Oligonucleotide binds to the ssDNA of $\Phi$ X174 DNA at the <i>attII</i> site, used for radio-labelling the $\Phi$ X174DNA during gel retardation assays
SWN40	GGA GAA ACG GGA ACA GGA GCG TCC ATG ACC ATA TCC GC	Primer pair used in Quickchange <sup>TM</sup> mutagenesis of <i>A. aeolicus recN</i> , generating the K35A mutation
SWN41	GCG GAT ATG GTC ATG GAC GCT CCT GTT CCC GTT TCT CC	
SWN42	GGG GAA ACT GGT GCT GGA GCA TCT ATC GCC ATT GAT GC	Primer pair used in Quickchange <sup>TM</sup> mutagenesis of <i>H. influenzae recN</i> , generating the K35A mutation.
SWN43	GCA TCA ATG GCG ATA GAT GCT CCA GCA CCA GTT TCC CC	
SWN44	GCG AGA CCG GCG GTG CAT CTA TTG CAA TAG ATG CC	Primer pair used in Quickchange <sup>TM</sup> mutagenesis of <i>E. coli recN</i> , generating the K35A mutation.
SWN45	GGC ATC TAT TGC AAT AGA TGC ACC CGC GCC GGT CTC GC	

## **2.2 Biological methods**

### **2.2.1 Propagation and maintenance of bacterial strains**

It was standard practice to grow primary overnight cultures in 5ml LB broth inoculated from single colonies on LB plates. Appropriate antibiotics were added as required. Overnight cultures were routinely grown at 37°C, unless stated otherwise, with gentle aeration in a tube rotator, and stored for up to 2 weeks at 4°C. Sterile glycerol was added to cultures to a final concentration of 30% for long term storage at –20°C. Secondary cultures were typically grown in 8ml Mu broth inoculated with 0.04–0.2ml overnight culture. Incubation was performed with vigorous aeration in a shaking water bath (Grants Instruments) at 37°C, unless stated otherwise, to the required optical density (OD) as measured at 650nm using a Thermo Spectronic 20 spectrophotometer or a Beckman Coulter DU530 spectrophotometer. An OD<sub>650</sub> of 0.48 was taken to be equivalent to  $3 \times 10^8$  cells/ml. All cultures were disposed of in 0.5% Trigene.

### **2.2.2 Harvesting bacterial cells from liquid culture**

Cells were harvested from culture by centrifugation. Volumes up to 2ml were harvested in a bench top microfuge at room temperature for 1 min at 13,000 rpm. Volumes up to 8 ml were harvested, at 4°C, in a Sorvall SS-34 rotor for 6min at 6,000rpm. While larger volumes 50 – 200ml were harvested at 4°C in a Sorvall SLA-1500 rotor at 6,000rpm and anything larger than this harvested by centrifugation in a Sorvall SLA-3000.

### **2.2.3 Preparation of bacteriophage P1 lysates**

P1 lysates for use in P1<sub>vir</sub>-mediated transductions were generated either in liquid culture, or on agar plates. To prepare lysates in liquid culture 8ml Mu broth was inoculated with between 0.2–0.5ml host strain overnight culture, and incubated at 37°C in a shaking water bath to an OD<sub>650</sub> of 0.3–0.4. Cultures were supplemented with 0.1 ml 0.5 M CaCl<sub>2</sub> and incubated for a further 10min, and then inoculated with approximately  $1 \times 10^7$ – $1 \times 10^8$  plaque forming units (pfu) of P1<sub>vir</sub> bacteriophage propagated in the wild-type strain W3110. The culture was returned to the shaking water bath until lysed, after which 0.5ml

chloroform was added and the cell debris pelleted in a refrigerated centrifuge at 10,000rpm for 15min. The supernatant was decanted, and 0.5ml of chloroform added prior to storage at 4°C.

Alternatively P1 lysates were generated on plates by incubating 8ml Mu broth (supplemented with 0.1ml 0.5M CaCl<sub>2</sub>), inoculated with between 0.2–0.5ml fresh host overnight culture, to an OD<sub>650</sub> of between 0.8–1.0. Fresh P1 agar plates were overlaid with 0.1ml of culture and 3ml 0.4% Mu soft agar, containing approximately  $2 \times 10^7$ ,  $1 \times 10^7$  and  $5 \times 10^6$  pfu of P1 phage propagated on strain W3110, and incubated for <18h at 37°C. Plates were examined for lysis and phage harvested from overlays in a Sorvall SS-34 rotor 4°C, with 1ml MC buffer, and 0.5ml chloroform, for 15min at 10,000rpm. The supernatant was decanted, and an additional 0.5ml chloroform was added for storage at 4°C.

#### **2.2.4 Transduction with Bacteriophage P1<sub>vir</sub>**

The transduction of nutritional markers was completed essentially according to the method of Miller (1972). Simply put, the recipient strain was grown to an OD<sub>650</sub> of 0.8, the cells collected, and resuspended in 2 ml MC buffer. Phage was added at a ratio of approximately 0.1-0.5:1 phage:bacterial cell, allowed to adsorb for 25-30 min at 37°C before addition of 0.2ml 1M sodium citrate to chelate the calcium, stopping infection. The transduction mixture was mixed with 3ml molten water agar (at 42°C), and overlaid onto the appropriate media. Incubation was allowed to proceed until transductant colonies were sufficiently developed, typically 24-48h.

The transduction of antibiotic resistance markers was essentially the same as for the transduction of nutritional markers although the ratio of phage to bacteria was 0.5-1:1. The transduction mixture was added to 3.5ml of molten 0.6% Mu agar and used to overlay the appropriate selective media. Incubation periods were generally 16-40h.

Bacteriophage P1 transductions were used to construct new genotypes. A P1 lysate of the donor strain was prepared by growing a culture of the required strain to an OD<sub>650</sub> of 0.4 in Mu broth containing 6.25mM CaCl<sub>2</sub>. The culture was then infected with  $1 \times 10^8$  pfu and incubated until cell lysis was

observed. Cell debris was removed by centrifugation, and the phage stored over chloroform.

### **2.2.5 Bacterial Transformation**

*E. coli* were routinely made competent for the uptake of plasmid DNA by the CaCl<sub>2</sub>-heat shock method. The required strain was grown to an OD<sub>650</sub> of 0.6-0.8 in LB broth, the cells harvested, and resuspended in 5ml of ice-cold 0.1M CaCl<sub>2</sub>. The cells were incubated on ice for 45min before being harvested and resuspended in 1ml ice-cold 0.1M CaCl<sub>2</sub> and incubated for a further 15min on ice. Typically, 0.5-1µg of plasmid DNA was mixed with 50-100µl of competent cells, and incubated on ice for 30 min. Heat shock was at 42°C for 2min, followed by a further 2min on ice. 1ml of SOC broth was added, and the transformation mixture incubated at 37°C for 45min (or 32°C for temperature sensitive strains/plasmids). The transformed cells were collected by centrifugation, resuspended in 100µl fresh SOC broth and spread onto LB agar containing the appropriate selection.

Chemically competent cells were also made in advance and stored at -80°C by growing a 300ml culture in SOC broth to an OD<sub>650</sub> of 0.4 – 0.6. Cells were then harvested at 4°C by centrifugation and resuspended in 90ml of buffer TFB-1 before incubation on ice for one hour. The cells were then collected once again by centrifugation and resuspended in 10ml of buffer TFB-2 before being aliquoted and stored at -80°C, such cells were competent for several months and could be used by thawed on ice and then treated as above, with the cells being mixed with DNA and heat shocked.

When higher transformation efficiencies were required, e.g. introducing linear DNA into a strain, the cells were electroporated. The required strain was grown to an OD<sub>650</sub> of 0.6 in 10ml SOC broth. The cells were harvested, and washed three times with 1ml of ice-cold 1mM HEPES buffer pH7.0. After the third wash, the cell pellet was resuspended in 100µl of ice-cold sterile water. Typically 0.1-0.5µg of donor DNA was mixed with 50µl competent cells before being added to a pre-cooled, 0.1cm electroporation cuvette. Electroporation was performed using a BioRad Gene Pulsar set at 1.8 kV,

25 $\mu$ F with Pulse controlled at 200ohms. The electroporated cells were immediately diluted with 1ml SOC broth, incubated for 1h at the appropriate temperature before being plated onto LB agar containing the appropriate selection for the transformation.

#### **2.2.6 Testing strain genotypes**

When it was necessary to test the phenotypes of several colonies, e.g. moving linked genes, the colonies were picked and streaked as a regular array (gridded) onto a master plate. Sterile velvets were used to print replicas of the master grid onto different media. Patches of the grid displaying the desired phenotype were purified from the master plate and retested. When it was necessary to test fewer colonies e.g. when moving a gene containing an antibiotic resistance marker as an insertion, the transductants were purified, and 10 $\mu$ l of the overnight culture tested directly.

#### **Confirmation of Genetic Markers**

Transposon and antibiotic resistance insertions were confirmed on the appropriate selective media. Auxotrophy was confirmed by streaking onto minimal media with and without the appropriate sugars and/or amino acids.

#### **Sensitivity to DNA Damage**

Semi-quantitative tests of DNA damage sensitivity were completed on LB agar with or without MMC at 0.2 $\mu$ g/ml and 0.5 $\mu$ g/ml. UV irradiation of a duplicate set was completed at a dose rate of 1 J/m<sup>2</sup>/sec, at a peak output of 245nm, for 30seconds (MMC containing plates) and 1min (LB agar plates).

Where appropriate, a quantitative measure of the ability of a culture to survive UV or  $\gamma$ -irradiation was made. Cultures were grown in LB broth to an OD<sub>650</sub> of 0.4, and the culture diluted by ten-fold increments in 56/2 salts. 10 $\mu$ l of the dilutions were spotted onto LB agar, and the spots allowed to dry before being exposed to the radiation. UV-irradiation was at 1 J/m<sup>2</sup>/sec for set intervals up to 60 seconds.  $\gamma$ -irradiation was from a caesium-137 source with an output of 923.3 Rads/min, for set intervals up to 48min. Irradiated plates

were incubated along with an unirradiated control for 16-24h before colonies of survivors were scored.

### **2.2.7 Transposon mutagenesis**

Mutagenesis was conducted using a mini Tn10 carried by  $\lambda$ NK1327 or by the EZ-system (Epicentre).

#### **Tn10-mediated transposon mutagenesis**

The  $\lambda$ NK1327 phage carries a mini-transposon, derived from Tn10, it only has a lysogenic life cycle, due to amber mutations, which inactivate genes essential to the lytic life cycle. Due to this it can randomly introduce kanamycin resistance genes into the chromosome, flanked by short repeats of IS10 sequences but not lyse the cell (Kleckner *et al.* 1991). It was utilised to mutagenise various strains in the synthetic lethal screen. Phage  $\lambda$ NK1327 was purified to a single plaque and then used to grow a plate lysate. Strain N1710 was grown either in TB broth overnight, or fresh in LB to an OD<sub>650</sub> of 0.4-0.6. The cells were then harvested by centrifugation and resuspended in 0.5 volumes of 10mM MgSO<sub>4</sub> before being incubated for 30min at 30°C to starve the cells. A 100 $\mu$ l aliquot of starved cells was mixed with 2.5ml of 0.7% TB overlay agar supplemented with 100 $\mu$ l of 10% (w/v) MgSO<sub>4</sub> and overlaid onto TB agar plates and once hardened the  $\lambda$ NK1327 stock was streaked onto the overlay. Plates were incubated overnight at 37°C and a single plaque was then removed as an agar plug and placed in 1ml of  $\lambda$  buffer and a drop of chloroform added to kill cells, 10 $\mu$ l of this was then spotted onto a fresh overlay. This gave a larger plaque, which was extracted and treated as the individual plaque to give the initial phage stock. Aliquots, (100-200 $\mu$ l) of this stock, were then mixed with a similar volume of the starved cells, incubated at 32°C for 10min, before being overlaid on to TB plates by mixing with 2.5ml of TB 0.7% agar. These overlays were sealed in a damp container (a “wet box”) and incubated for 6-8h or until lysis occurs and plates appear clear. Once lysis occurred the plates were chilled to 4°C and then flooded with 5ml of pre-chilled  $\lambda$  buffer. These were left overnight to let the phage leech out into the

buffer, which was then collected. To the collected buffer 0.5ml of chloroform was added to kill cells, cell debris was removed by centrifugation (10,500rpm, 20min, SS-34 rotor). The supernatant was decanted to a fresh tube and a few drops of chloroform added, before being stored at 4°C.

The protocol used for mutagenesis is essentially that described by Kleckner *et al* (1991). A 50ml culture of the strain to be mutagenised is grown in LB broth, supplemented with 0.2% MgSO<sub>4</sub> and any antibiotic selection required, to an OD<sub>650</sub> of 0.3-0.4. The cells were harvested by centrifugation and resuspended in 2ml of 10mM MgSO<sub>4</sub>, 200µl aliquots of cells were then added to 200µl aliquots of λNK1327 either, undiluted or diluted to 10<sup>-1</sup> or 10<sup>-2</sup>. These were then incubated for 15min at 37°C with vigorous shaking in a water bath, to allow phage adsorption. Once incubated 200µl of 1M Sodium citrate was added to each dilution to chelate calcium ions and so halt infection. The cells were then allowed to recover at 42°C for 1h in LB broth, before being harvested by centrifugation and resuspended in 500µl of 56/2:Sodium citrate (1:1). These cells were then plated onto LB media supplemented with kanamycin, and any other supplements required by the strain. The plates were then incubated for 16h at 37°C and the cells collected by re-suspension of colonies in LB broth. Library size i.e. the number of inserts was estimated from the number of colonies present on the 10<sup>-1</sup> and 10<sup>-2</sup> dilution plates. For storage glycerol was added to a final concentration of 30% (v/v) glycerol.

### **EZ::TN transposome mutagenesis**

The EZ::TN <KAN-2> Tnp Transposome kit<sup>TM</sup> (Epicentre) was used as an alternative to Tn10 mutagenesis based on λNK1327. The method involves introducing a transposome, essentially a transposon and its specific transposase protein into a cell. The transposase is rapidly degraded and so the insertion is not mobile. The method used was similar to that described in the Epicentre user's guide. Cells were made electro competent by growing the desired culture in 50ml LB broth to an OD<sub>650</sub> of 0.3-0.4. They were then incubated on ice for 15min before being harvested by centrifugation and resuspended in 20ml, of ice cold, 10% glycerol. This was repeated twice and the cells then



resuspended in just 200µl of 10% glycerol. These were kept on ice, and then 40µl aliquots placed into pre-cooled electroporation cuvettes and gently mixed with 0.25µl of the supplied transposome solution. The mixture was then electroporated (2kV/cm, 25µFD, 200ohms with a Time constant of 4.0-4.8), 1ml of SOC broth was immediately added to the electroporated cells and the cells were then incubated at 30°C for one hour to recover. The cells were then diluted 5-fold and plated onto LB agar supplemented with Kanamycin to select for the insert and grown for 16h at 37°C and the cells collected by re-suspension of colonies in LB broth.

### **2.2.8 Measures of cell survival following chromosome cleavage by I-SceI**

#### **Qualitative Viability Assay**

To assay survival after I-SceI expression, 10µl of fresh overnight culture of strains carrying both an I-SceI cleavage site and the chromosomal arabinose inducible I-SceI expression cassette were streaked on LB agar plates supplemented with, or without, arabinose at a final concentration of 0.2% (w/v). Growth was scored after overnight incubation at 37°C.

#### **Quantitative Viability assay**

For quantitative analysis, survival was measured after transient I-SceI expression in liquid culture, 40µl of fresh overnight culture was diluted into 8ml Mu broth (approximately a 200-fold dilution), and grown at 37°C to an OD<sub>650</sub> of 0.2 (unless otherwise stated). Cultures were then split in two, with glucose added to one half, and arabinose to the other, to final concentrations of 0.2% (w/v). Incubation was continued for 30min, unless stated otherwise, before cultures were serially diluted in 56/2 salts. 10µl aliquots were spotted onto LB agar and scored for growth after overnight incubation. A relative viability was calculated from a comparison of glucose and arabinose grown cultures (Meddows *et al.* 2004). Data from all lab members experiments on I-SceI viabilities were pooled and viabilities from all experiments incorporated into the data set.

## **2.3 DNA preparation and analysis**

### **2.3.1 Preparation of Plasmid DNA**

Plasmid DNA was extracted from overnight cultures using the appropriate Qiagen kit QIAprep Spin Miniprep Kit (1.5-3ml of culture), or the QIAGEN-tip 100 (60-100ml of culture) following the manufactures instructions. Briefly, the cells were harvested by centrifugation, and resuspended in 50mM Tris-HCl pH8.0, 10mM EDTA, 100µg/ml RNase A. An equal volume of lysis solution (200mM NaOH, 1% SDS) was added to the cell suspension, incubated for 1-5 min at room temperature, and neutralised with 1.4 volumes of 3M potassium acetate, pH5.5. The cell debris was removed by centrifugation and when purifying DNA on a small scale, the clear lysate was added to a QIAprep spin column, allowing plasmid DNA to bind the silica-gel membrane in the column. The DNA was then washed with an ethanol based buffer (Buffer PE, Qiagen), before being eluted in 10mM Tris-HCl pH8.5. The principle of the larger scale plasmid purification is the same, although QIAGEN-tip 100s, and different buffers were used. The column matrix was pre-equilibrated with 750mM NaCl, 50mM MOPS, pH7.0, 15% isopropanol, 0.15% Triton X-100 before plasmid DNA was bound. The column was washed with the same buffer but containing 1M NaCl, and the DNA eluted in 1.25M NaCl, 50mM Tris-HCl, pH8.5, 15% isopropanol. Eluted DNA was precipitated by addition 0.7 volumes of isopropanol and pelleted by centrifugation. The pellet was washed with 70% ethanol and the DNA resuspended in 50-200µl TE buffer, depending on the copy number of the plasmid.

### **2.3.2 Preparation of chromosomal DNA**

Chromosomal DNA was extracted from cells in one of two ways:

#### **Extraction using Phenol Chloroform**

Strains were grown to the required optical density, the cells (100ml) harvested by centrifugation and resuspended in 0.5ml TNE. 50µl each of 10% Triton X-100 and 5mg/ml lysozyme were added, and incubated at 37°C for 30min. 60µl of a 5mg/ml solution of proteinase K was then added, followed by incubation

at 65°C for 2h. A standard phenol-chloroform extraction was completed (Sambrook *et al.* 1989), the DNA precipitated with 0.1 volumes sodium acetate and 2 volumes of ethanol, and spooled out into 70% ethanol. The DNA was then allowed to dissolve overnight in 1ml of TE with gentle agitation at 4°C. This method produces about 1mg of chromosomal DNA.

### **Extraction using CTAB**

Strains were grown to the required optical density (OD<sub>650</sub> of 0.8 – 1) then harvested by centrifugation. The pellet was resuspended in 6ml CTAB extraction solution and 0.2mg/ml final concentration of proteinase K and 80µg/ml RNase H added, followed by incubation at 37°C for 1h. 4ml of chloroform:isoamylalcohol (24:1) was added and the sample emulsified by vortexing. This was centrifuged (15min, 12,000rpm in SS-34 rotor) and the upper layer removed and mixed gently with 18ml of CTAB dilution solution until a precipitate appeared. This was allowed to settle to form a pellet and the supernatant removed, before rinsing the precipitate with 25ml of 0.4M NaCl in TE, this was decanted and the pellet washed a second time. The pellet was then mixed with 5ml of 1.42M NaCl in TE and gently mixed until the pellet becomes transparent. The mixture was then ethanol precipitated by addition of 10ml of 70% ethanol, this was decanted and the precipitate washed with 10ml of 70% ethanol a second time. The resultant pellet was resuspended in 4.5ml TE and precipitated using 4ml of 3M sodium acetate:isopropanol (1:7). The precipitant was pelleted by centrifugation and washed twice with 70% ethanol, which was completely removed and the pellet dissolved overnight in 300 - 500µl of TE and giving 0.5 – 1mg of chromosomal DNA.

### **2.3.3 Agarose gel electrophoresis**

DNA fragments were routinely separated with 1% Seakem agarose (Flowgen) although 0.8%-1.2% gels were used to separate larger or smaller fragments. Gels were made using TBE, and stained with 0.2µg/ml ethidium bromide before casting. Before loading the DNA sample was mixed with 0.2 volumes of Ficoll gel loading buffer. DNA was separated on the gel at a constant

voltage of 5V/cm was used, and fragment size was determined using 0.25µg of 1kb ladder (NEB). DNA was visualised using UV light, and documented with a BIO-RAD Gel Doc EQ and the Quantity one 1-D analysis software.

DNA fragments separated on agarose were excised using a clean scalpel blade, and the agarose removed using the QIAquick Gel Extraction kit (Qiagen). The agarose containing DNA was melted at 55°C in a solubilisation buffer. This solution containing the DNA was passed through a QIAquick spin column, selectively binding the DNA. The DNA washed on the column using an ethanol based buffer (buffer PE, Qiagen), and eluted using 35-50µl Tris-HCl pH8.5. In some cases SyBr Green was added to the DNA (approximately 2000-fold dilution) before loading and the DNA visualised using blue light, to prevent UV induced damage of the excised DNA. The DNA was excised and purified as described.

### **2.3.4 Cloning DNA fragments**

#### **DNA Restriction**

A standard restriction digest was completed using 1-2.5µg DNA, a buffer appropriate to the enzyme (supplied by the manufacturer), and 5-10units of enzyme in a 20 or 50µl final volume. Incubation times varied from 1h to overnight. The DNA was purified from the reaction using gel purification (Qiagen).

#### **Ligation**

Appropriately prepared DNA fragments were ligated using T4 DNA ligase, and the supplied buffer. The fragments to be ligated were quantified, and mixed in a ratio so that the number of DNA ends of the insert were in a three-fold excess to the number of ends of the vector. A typical reaction was completed in a 20µl final volume usually cohesive-end ligations were incubated at 15°C for 4h to overnight, and blunt-end ligations were incubated at 15°C overnight. 5µl of the ligation mix was transformed into competent DH5α, and spread onto LB agar containing the appropriate selection.

### **Polymerase chain reaction (PCR)**

DNA amplification was completed using the Polymerase Chain Reaction (PCR) according to published methods (Mullis and Faloona 1987; Saiki *et al.* 1988), and following the recommendations of the enzyme manufacturer. *Taq* DNA polymerase was used. However, in most instances the higher fidelity Dynazyme or Phusion polymerases (NEB) were used for preparing DNA for sequencing, and screening for chromosomal deletions/insertions, as well as the cloning of genes by PCR. When a bacterial colony was to be used as the source of template DNA, a single colony was transferred to 50µl sterile distilled water using a sterile toothpick. The suspension was vortexed for 3min, and incubated at 37°C for 15min; 1µl was used in the PCR reaction.

### **QuikChange™ Mutagenesis**

Point mutations were introduced into DNA using the QuikChange™ protocol (Stratagene). Two complementary primers containing the mutation were used to amplify a circular template. The template DNA was degraded by *DpnI* digestion, and the reaction transformed into DH5α. Positive clones were screened for by DNA sequencing, or if a novel restriction site was created, or an existing one removed, restriction analysis.

### **2.3.5 Automated DNA sequencing**

The target sequence was amplified from the template using standard PCR protocol and Phusion DNA polymerase. The product of the reaction was gel purified into 50µl of dH<sub>2</sub>O, and used as a template for cycle sequencing. Cycle sequencing used 4µl each of BigDye (ABI Prism) and HalfTerm (Genpack), and 0.3µl of primer at 50nM concentration in a 20µl reaction volume. The reaction was cycled 25 times at 96°C for 30seconds, 50°C for 15seconds and 60°C for 4min. The reaction products were purified by precipitation with 0.1 volumes 3M sodium acetate and 2.5 volumes 100% ethanol, and analysed in an ABI model 373A automated sequencer. Good quality template DNA in the cycle sequencing reaction allowed DNA sequence to be read up to 1000bp

from the primer. Sequencing was performed by the BSAU, University of Nottingham, or MWG.

### **2.3.6 Creating gene deletions - Recombineering**

Gene deletions were created according to the method described by Datsenko and Wanner (2000). Briefly, a PCR product was created so as to insert an antibiotic resistance gene with flanking homology that complements the ends of the region to be deleted. MG1655 cells were transformed, at 30°C, with pKD46 selecting for the ampicillin resistance. Cultures of this were grown in 8ml of Mu broth supplemented with 0.15% (w/v) arabinose, and ampicillin at 30°C. Once at an OD<sub>650</sub> of 0.5 cultures were immediately cooled on ice for 10min. Cells were harvested by centrifugation and washed three times with 4ml of 1mM HEPES pH7.5, before finally being resuspended in 40µl of 1mM HEPES pH7.5. All 40µl of cells were transformed by electroporation at 1.75 kV using between 10–100ng chilled PCR product, eluted from a Qiagen column in 1mM HEPES pH7.5. Cells were resuspended in 3ml SOC, and allowed to recover at 30°C for 1h. Cells were harvested by centrifugation, resuspended in 50µl SOC, and incubated overnight, at 37°C, on LB agar supplemented with the appropriate antibiotic. Single colonies were inoculated in regular arrays on selection medium and replica plated onto LB agar supplemented with ampicillin to confirm the spontaneous loss of pKD46.

## **2.4 Protein preparation and analysis**

### **2.4.1 SDS-PAGE**

Proteins were separated by one-dimensional electrophoresis using the BioRad Mini-Protean II gel apparatus throughout. The protein gels were separated into resolving, and stacking components. The resolving gel contained 12.5% acrylamide/ bis-acrylamide 29:1, 0.375M Tris-HCl pH8.8, and 0.1% (w/v) SDS. The stacking gel contained 5% acrylamide:bis-acrylamide 29:1, 0.125M Tris-HCl pH6.8, and 0.1% SDS. Both portions of the gel were polymerised using 0.08% ammonium persulphate and 0.08% tetramethylethylenediamine (TEMED). Protein samples were prepared by mixing at a 5:1 ratio of the

sample and 5x SDS loading buffer, they were then heated to 95°C for 2min. Molecular Weight Marker Standards (Bio-Rad or Fermentas PAGERULER™), and protein samples were loaded onto the gel, and run at 200V for 75 min in SDS PAGE running buffer. Gels were stained using PageBlue™ (Fermentas) as described by the manufacturer. Briefly the gel was washed twice with hot dH<sub>2</sub>O and then stained with 20ml of the PageBlue™ solution, heated in the microwave to approximately 90°C and allowed to stain for 15min. The gel was then de-stained by soaking in dH<sub>2</sub>O for 30min.

#### **2.4.2 Protein Overexpression**

Overexpression of cloned genes in *E. coli* was tested on a small scale before the volumes used in purification were inoculated. Samples from fresh 5 ml overnight cultures, made from freshly transformed colonies, were used to inoculate 8 ml of Mu broth; the appropriate selection was maintained at all stages. The secondary cultures were grown to an OD<sub>650</sub> of 0.4-0.5 at 37°C and a sample of the culture was removed as a zero time point, the remainder split into two, and inducer added to half. Unless otherwise stated in the text, induction was with 1mM IPTG for the *lac* or T7 promoter driven systems and 0.5% (v/v) arabinose for the P<sub>BAD</sub> promoter systems. Other promoter systems used are described in the text. Overexpression was typically for 3-4 hours, although a total cell protein was analysed by SDS-PAGE, using 10µl of 0.5ml culture collected by centrifugation and lysed with 100µl 1x SDS-PAGE running buffer.

Overexpression on a preparative scale was performed using 1l baffled flasks containing 400 ml Mu or YT broth inoculated with 3-5ml of a fresh culture. Details are otherwise as for small-scale overexpression with large scale overexpression up to 10l performed in a Fermenter 310/360 fermenter system (Electrolab). The fermenter and the growth vessel with media and associated apparatus was autoclaved at (123°C, 90min) to sterilize. Once sterilized the growth vessel was reconnected to the external fermenter components, which provides both a jacket for heating and a sealed water cooling system for the growth vessel. Aeration of the media was with a filtered air supply at 6-

7litres/min flow rate. The apparatus and media was left to equilibrate overnight to 37°C. The following day the media was checked for bacterial contamination and the temperature. The dissolved oxygen probe was calibrated so that the culture at full aeration with rapid mixing from the rotor (300-400rpm) was 100% aeration. At this point anti-foam was added if required and any antibiotics needed for selection. Introduction of cell starter culture, antifoam and antibiotics was via self-sealing membrane to prevent contamination. Inoculations were at a 1:100 (occasionally 1:50 if the strain is slow growing) ratio of starter culture to growth media. The culture was grown to the desired OD<sub>650</sub>, with rotor speed regulated by the dissolved oxygen levels to ensure 100% aeration. Temperature was controlled (usually at 37°C), although pH was not. In all cases cells were harvested by centrifugation and usually stored as cell pellets at -80°C.

#### **2.4.3 Isolating soluble proteins**

Cell pellets to be lysed were removed from -80°C storage, thawed in 0.01 volumes of lysis buffer (as page 41). Lysis was performed by sonication of the cells, using a Soniprep150 Ultrasonic disintegrator (Sanyo Gallenkamp) with either a 9.5mm probe or an exponential micropobe depending on the cell volume. Cells were sonicated on ice 3-5 times for 5-10 seconds with a microprobe, for volumes below 2ml. For larger volumes the cells were sonicated 3-5 time for 20-30 seconds, using the 9.5mm probe.

After lysis cell debris and soluble proteins are separated by centrifugation, 10,000g, for 30min, at 4°C. Depending on volume this was done either in a benchtop microfuge, or larger volumes processed using a Sorvall, SS-34 rotor.

#### **2.4.4 Ammonium sulphate precipitation**

Ammonium sulphate precipitations were performed by addition of ammonium sulphate to cell extracts to the desired percentage of saturation. This was done slowly to ensure all the ammonium sulphate dissolved and then samples allowed to precipitate for 1h, usually at 4°C, with continuous mixing. The precipitated material was harvested by centrifugation (10,000g for 30min).



#### **2.4.5 ÄKTA operation**

Protein purifications were performed using ÄKTA FPLC system (GE Healthcare). All of the columns used were purchased pre-packed from GE healthcare, unless stated otherwise. As a general rule all elutions were performed in a 20mM Tris-HCl pH7.5, 1mM EDTA based buffer (buffer A) with NaCl gradients used to elute proteins. All the buffers were vacuum filtered through a 0.2µm membrane to ensure particle removal and to degas the solution. Fractions were collected and those corresponding to the detected UV peaks were analysed by SDS-PAGE to determine where the protein being purified was located. Such fractions were pooled and then further purified. For details of RecN purification see Chapter 5.

#### **2.4.6 Determination of protein concentration**

Protein concentrations were determined using the BioRad Protein Assay kit, a modification of the Bradford Assay (Bradford 1976). The OD<sub>595</sub> of a series of 1-24µg/ml aliquots of BSA protein standard were determined, and compared to the OD<sub>595</sub> of the sample protein.

#### **2.4.7 Measuring ATP hydrolysis,**

ATP hydrolysis was measured by quantifying the production of inorganic phosphate, which was measured by monitoring the change in absorbance at 660nm of a mixture of malachite green and ammonium molybdate (Bird *et al.* 1997). ATPase activity was either assayed as a direct measure of inorganic phosphate produced over 30min or as a time course where samples of the reaction were taken at set time points. The proteins were incubated at 37°C (or 55°C for *A. aeolicus* RecN) in 1x helicase buffer, with the addition of 5mM MgCl<sub>2</sub> and 5mM ATP (unless otherwise described). In certain reactions ΦX174, pBR322 duplex plasmid DNA, or linear single stranded or duplex DNA oligonucleotides of 60, bases or base pairs respectively, were used. At the set time points, 100µl samples of the reaction mixture were taken and added to 800µl of a 3:1 mixture of 0.045% malachite green: 4.2% ammonium molybdate in 4M HCl, incubated for 2min at room temperature, and 100µl of a

34% sodium citrate solution added. The colour was allowed to develop for 30min and the absorbance was read at 660nm (using a Beckman Coulter DU530 spectrophotometer), and compared to a standard curve constructed using known concentrations of  $\text{KH}_2\text{PO}_4$ .

#### **2.4.8 Glutaraldehyde cross-linking**

To perform cross-linking, 10 $\mu\text{l}$  samples of the protein to be cross-linked, at 1mg/ml concentration were aliquoted into separate tubes and incubated with an equal volume of glutaraldehyde solution at 0.01-2% final concentration, for 15min at room temperature. Reactions were stopped by addition of an equal volume of 0.5M Tris-HCl pH7.0. Samples were then analysed by SDS-PAGE.

#### **2.4.9 Electron microscopy**

Electron microscopy was performed with the assistance of Dr. Katy Evans (University of Nottingham). The apparatus used was a transmission electron microscope (TEM), specifically a JEOL JEM 1010 TEM at 100kV at magnifications ranging from 100 – 500k. Using formavar/carbon grids (Agar scientific) with a 200 Cu mesh. Proteins were stained with 1% PTA or 1% uracil acetate.

#### **2.4.10 Labelling of Oligonucleotides and gel retardation assays**

Oligonucleotides that were to be used in gel retardation assays were labelled using T4 kinase (NEB) to replace the gamma phosphate with radioactive [ $\alpha^{32}\text{P}$ ] ATP following the manufacturer's instructions. Unincorporated oligonucleotide were removed using Bio-Rad exclusion columns with a 10base pair cut off. The quantity of labelled DNA recovered was estimated at 95% and the final concentration of labelled oligonucleotide calculated by comparing the total radioactivity used to the final emittance value of the labelled oligonucleotides assuming 90% recovery of the DNA.

Oligonucleotides used to make duplex DNA substrates were annealed to each other by mixing cold:labelled DNA at a 3:1 ratio quantities of each, heating to 95°C for 5min and then allowing the mixture to cool slowly to room temperature. The mixture was then analysed by polyacrylamide gel

electrophoresis and the annealed product visualised and excised from the gel, with a clean scalpel blade, before being eluted from the gel slice by placing it in TE buffer overnight at 4°C.

Retardation assays were performed using a 10% acrylamide/bis-acrylamide 29:1 TAE gel, which once set was placed in the gel running tank, with two litres of 1xTAE running buffer and left to equilibrate at 4°C overnight. DNA binding reactions were conducted in a fluid volume of 20µl and contain the DNA substrate at 2nM, as well as the protein and other supplements as specified, although usually RecN reactions were performed in 1xHB (page 41) supplemented with final concentrations of 5mM MgCl<sub>2</sub>, 1mM nucleotide and protein concentrations of 0-500nM. The binding reactions were then loaded onto the gel and run at 160Volts/Cm for 75 min. Once run the gel was dried using an air dryer (BIO-RAD model 583) and then the presence of radiolabeled DNA visualised by exposure to a phosphoimage screen (Molecular dynamics) and scanned (Molecular dynamics Storm scanner 840).

# Chapter 3

## The essential nature of the OxyR regulon in the absence of recombination

To isolate the genes involved in DNA repair, studies have often made use of random mutation and the detection of phenotypes, specifically, deficiencies in the process being studied. This includes studies into DSB repair, the enzymology of which has been well characterised in *E. coli* (as described in Chapter 1). Clark and Margulies identified the first recombination deficient mutants in 1965, by making mutation libraries of F<sup>-</sup> cells and identifying those mutants which had lost the ability to acquire genetic markers from Hfr strains. These mutants could not utilise HR and were therefore referred to as *rec*<sup>-</sup>. The first mutant found, named *recA*, was also shown to be sensitive to DNA damaging agents, in particular those, which are known to cause DSBs (Clark and Margulies 1965). Other *rec*<sup>-</sup> mutants were discovered later that year, initially identified as being sensitive to X-rays. They were subsequently shown to have reduced recombination efficiencies and were named *recB* and *recC* (Emmerson and Howard-Flanders 1965). Many of the other *rec* genes were discovered in a similar manner, including *recN*. Unfortunately, this methodology has limitations as it can only be used to identify genes that, when mutated, cause a detectable phenotype, which is actually being screened for at this time. Secondly, the function lost cannot be essential; otherwise, the cells will simply die. Therefore, to identify novel genes involved in recombination, and in particular the repair of DSBs, a synthetic lethal screen, similar to that used in yeast was used (Bender and Pringle 1991).

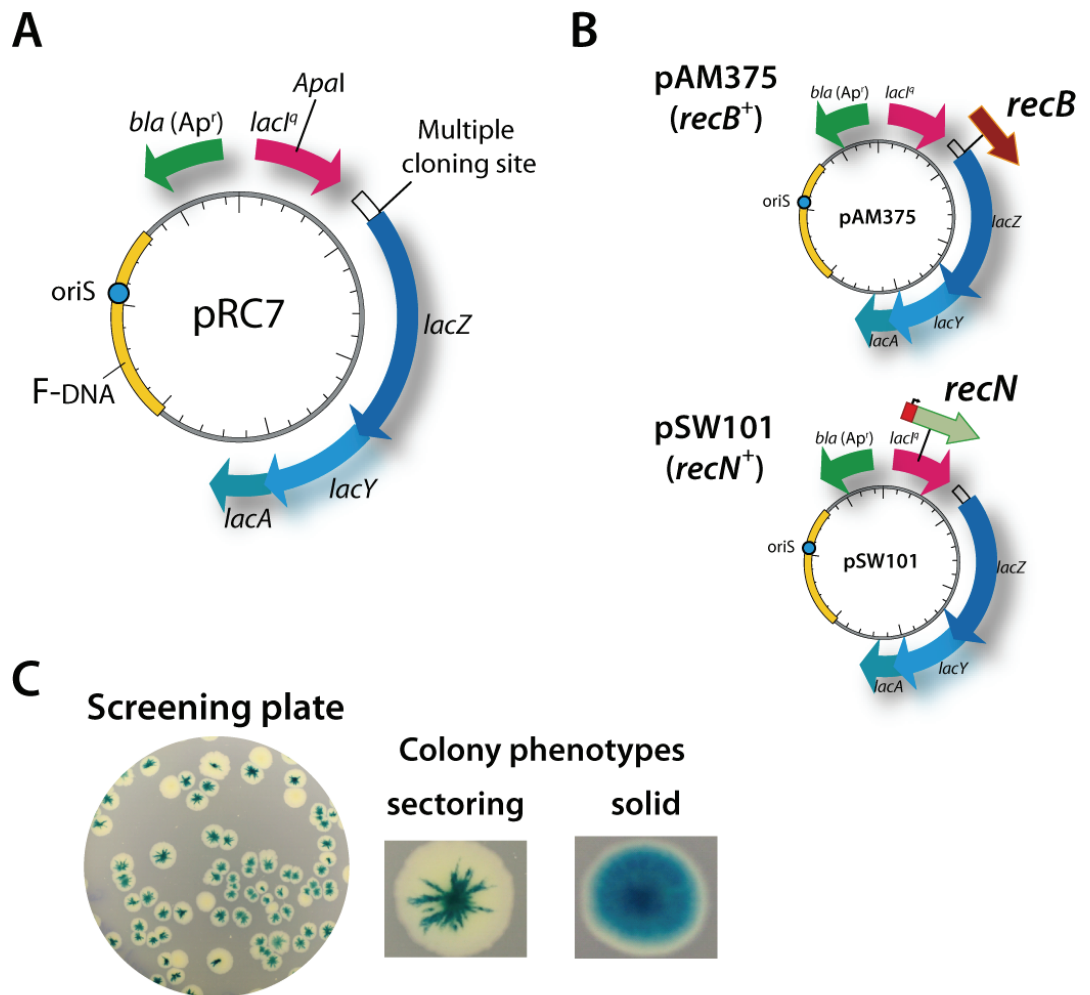
### 3.1 Identification of genes that are synthetically lethal with the recombination genes *recB* and *recN*

The role of the *recN* gene product in the repair of DSBs is currently ambiguous. It may act very early in repair, flagging a DSB, or more likely acts later to ensure high fidelity repair of the break (Meddows *et al.* 2005; Sanchez *et al.* 2006). The *recB* gene product's role is well defined, being one of the components of the RecBCD end processing complex that resects the ends of a break, allowing the initiation of homologous recombination (Spies *et al.* 2005). Therefore, my initial studies into double strand break repair began by examining the possibility that the *recB* and *recN* genes may have functions that have yet to be identified because these functions are masked by the overlapping activities of other gene products. To investigate this possibility I initiated a screen for mutations that are synthetically lethal with *recB* or *recN* null alleles. The analysis of such mutations might then provide insight into recombination and perhaps identify new cellular interactions that shed light on the role of RecN in particular.

Bernhardt and de Boer (2004) originally established the *E. coli* synthetic lethal screen adopted. The screen involves covering a chromosomal deletion of the gene to be studied with a functional copy of the gene borne on a derivative of plasmid pRC7. This is a low copy number, mini-F vector plasmid, which lacks the usual F-factor stabilisation system, and is therefore unstable (Figure 3.1 A) (Koop *et al.* 1987). The synthetic lethal screen relies on the unstable nature of this plasmid, as retention suggests a selection pressure acting to keep the plasmid in the cell. Derivatives of pRC7 were made which were either *recB*<sup>+</sup> (pAM375) or *recN*<sup>+</sup> (pSW101). The *recN* derivative included not only the coding sequence of the gene, but also its promoter to ensure the correct regulation of expression, while the *recB* gene expression was regulated by the *lac* promoter (Figure 3.1 B). For screening purposes the plasmid carries the *lacZYA* reporter operon. Loss of the plasmid can therefore be detected in a  $\Delta lac$  background by the segregation of *lac*<sup>-</sup> clones. On media containing the  $\beta$ -galactosidase indicator, X-gal, these *lac*<sup>-</sup> clones appear as white colonies or white sectors within a blue colony. Synthetic lethality between the primary

mutation, which is covered by a pRC7 derivative carrying the gene, and a secondary mutation is revealed when plasmid free clones fail to grow, resulting in uniformly blue colonies (Figure 3.1 C).

The two pRC7 derivatives were transformed into a  $\Delta lacZ$  background, which was also deficient in their respective gene, making activity dependent on the plasmid-borne copy (*recN* strain is SW1054 and *recB* strain AM1581). These strains all had a sectoring phenotype, indicating plasmid loss, on X-gal (Figure 3.1 C). The relevant plasmid-borne gene was able to complement for the known UV and mitomycin C sensitivity of a *recB* mutant and the mitomycin C sensitivity of a *recN* mutation. Mutant libraries were then made of these strains, using either the EZ::TN <KAN-2> Tnp Transposome kit<sup>TM</sup> (Epicentre), or phage  $\lambda$ NK1327 (Kleckner *et al.* 1991), as described in Chapter 2. Both methods randomly insert kan resistance genes into the genome. Screening was performed by plating the libraries onto LB agar plates supplemented with X-gal (6 $\mu$ g/ml), and IPTG (5mM). Colonies were allowed to grow at 37°C for 24-48 hours. Any colonies showing a solid blue phenotype, indicating retention of the plasmid, were purified onto fresh agar plates containing X-gal and IPTG (Figure 3.1 C). If they still had the solid blue phenotype, a stock of phage P1 was grown on the mutant isolate and the phenotype confirmed by transduction of the kanamycin resistance marker into the parental strain from which the mutant libraries were made. This was necessary as a relatively large number of colonies were actually shown to be false positives. During the screening, over 60 insertions in the *recB* screen and 23 insertions in the *recN* screen were shown to be false positives (Table 3.1). In these instances the solid blue colony phenotype was not due to plasmid retention because of the gene it carried, but some other factor. Commonly, this was due to integration of the plasmid into the chromosome, although in some instances the mutations were found to have increased the stability of the plasmid. Once the synthetic lethal phenotype was confirmed, the location of kanamycin insertion was determined by sequencing.



**Figure 3.1. Plasmid constructs used in the synthetic lethal screen.**

- A)** Plasmid map of the unstable mini-F plasmid pRC7, which was used in this study.
- B)** Maps of the plasmid constructs derived from pRC7 containing either the *recB* or the *recN* gene coding sequence, which was cloned into them as shown by the additional arrows. The *recN* gene also had its promoter cloned with it (both constructs are described in Table 2.4).
- C)** Screening for synthetic lethal mutations. Colonies were allowed to grow on X-gal containing media, resulting in either a sectoring or a solid blue colony phenotype.

Despite screening several *recN* libraries, which included over 12,000 separate insertion events, no synthetic lethal mutations were identified (table 3.1). Although solid blue colonies were present, none of them could be reconstructed by a backcross (P1 transduction of the kan marker into the

parental strain). It would seem that there is no mutation, at least of a non-essential gene, that is synthetically lethal with *recN*. This does not rule out the possibility of RecN protein having interaction or complex forming partners, but if they exist they are either essential to cell survival, or their absence is not lethal, even in the absence of RecN. However, I have to accept the possibility that mutations that cause synthetic lethality do exist and could be identified by transposon mutagenesis, but were missed by my screen. There are several thousand non-essential genes in *E. coli*. Although I screened 12,000 insertion events in total, the possibility of missing an insertion into one gene still exists, especially if the gene were to be smaller than average.

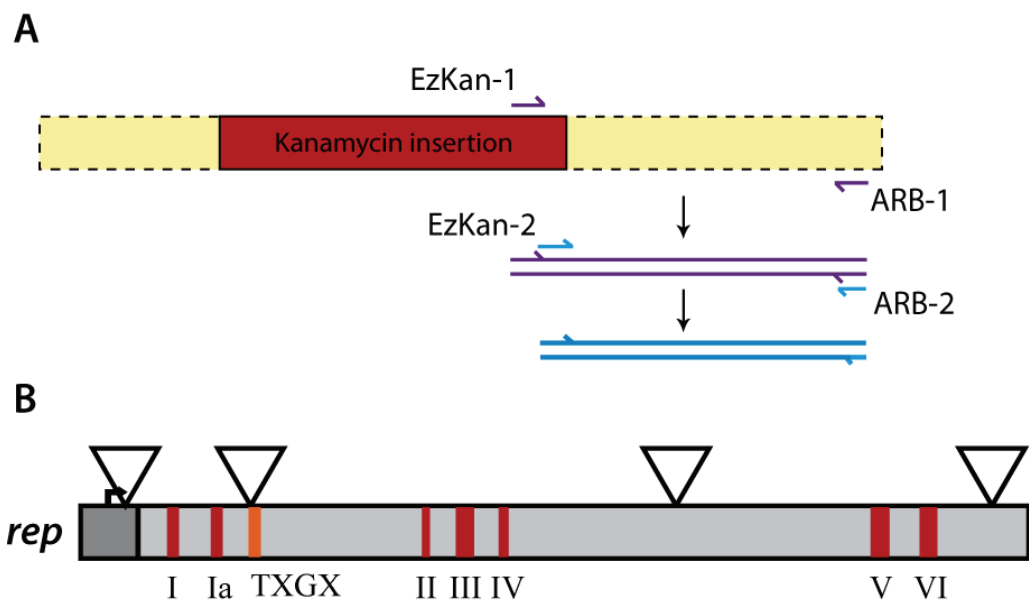
In the case of the *recB* screen, seven synthetic lethal mutants were identified, all of which were located inside the *rep* gene (Figure 3.2). Two of the mutations were confirmed by sequencing. This was achieved, by colony PCR, using a set primer at the 3' end of the kanamycin resistance gene (EzKan-1) and an arbitrary primer (ARB-1) that will bind randomly within the *E. coli* genome. The product of this reaction was amplified if necessary by a secondary, nested, PCR (using EzKan-2 and ARB-2 primer set) and then sequenced (Figure 3.2 A). The identity of the other five mutations as *rep* insertions was confirmed by the introduction of a functional *rep* gene into to the chromosome allowing normal growth (Table 3.1). This data was supplemented by work by Dr. Akeel Mahdi (University of Nottingham), who also identified and sequenced two other *recB* synthetic lethal mutations, both of which were also in *rep* (Dr. Akeel Mahdi, personal communication) (Figure 3.2 B).



Primary mutation	Number of libraries made	Estimated number of inserts screened	Number of colonies screened	Number of potential synthetic lethals found	Confirmed synthetic lethal mutations
<i>recB</i>	7	9,200	~25,000	62	7 + 1 synthetically sick mutation
<i>recN</i>	14	12,450	~50-75,000	23	0

**Table 3.1** Table showing the scale of the synthetic lethal screens and number if mutations identified as interesting.

Inserts screened is an estimate of the number of inserts made, colonies screened is estimated on plate numbers assuming 50-100 colonies per plate on average



**Figure 3.2.** Locating mutations that are synthetically lethal with *recB*.

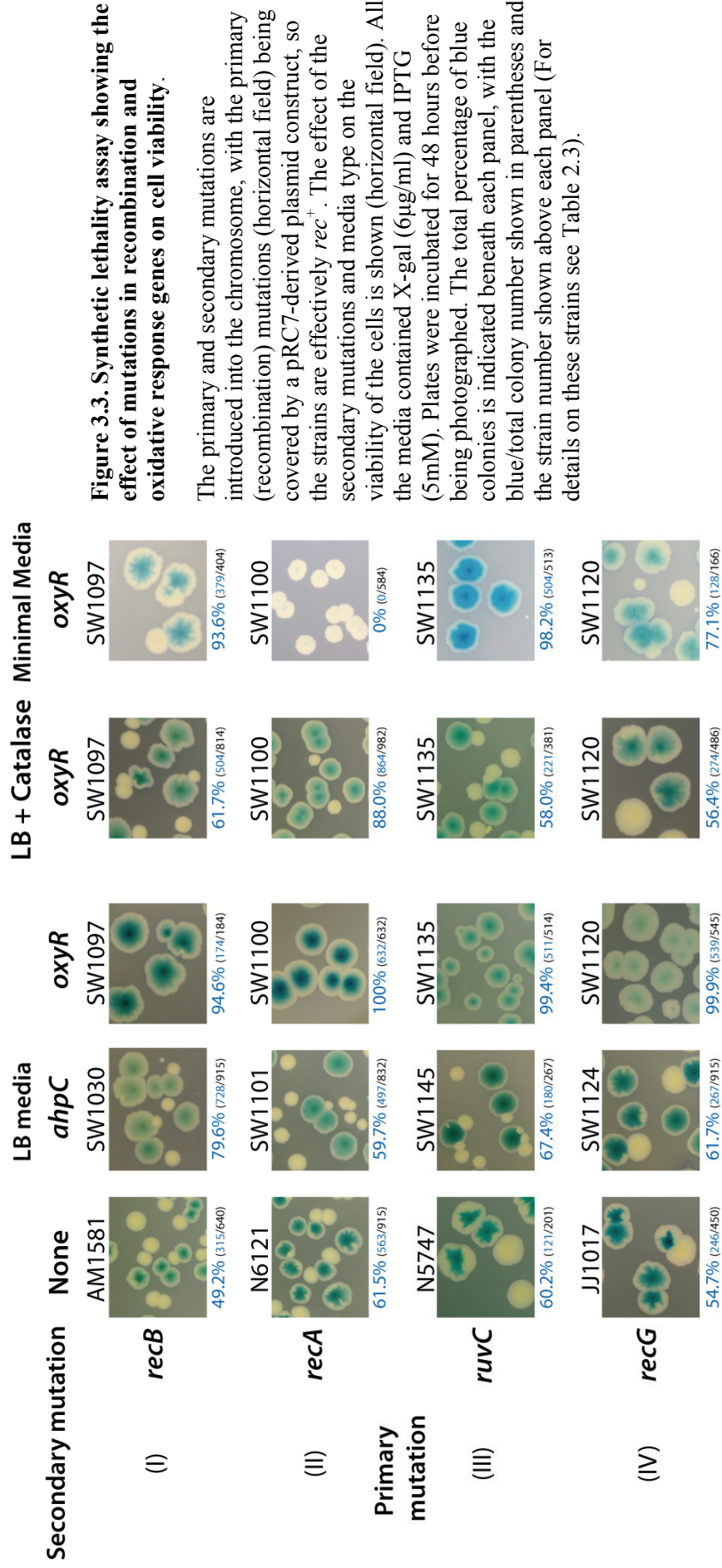
- A)** Schematic showing the PCR reactions used to isolate and sequence the DNA fragment showing the location of the insertion. The kanamycin insertion is shown along with a piece of chromosomal DNA. The binding positions of the primary primer pair (EzKan-1 and ARB-2) and the nested primer pair (EzKan-2 and ARB-2) are also shown (Primers are defined in Table 2.6).
- B)** Schematic of the *rep* gene modified from Korolev *et al* (1997), the coding sequence (light grey) and the promoter region (dark grey) are shown along with the conserved helicase (red) and TXGX motif (orange). The location of the four known inserts is also shown (triangles).

The synthetic lethality of *rep* and *recB* mutations has already been documented (Uzest *et al.* 1995), but the repeated identification of *rep* did demonstrate the robust nature of the system and would seem to confirm that the absence of *recN* synthetic lethal mutations is a real result and not a fault of the system. The high incidence of *rep* inserts could also suggest that the gene is perhaps in a recombination ‘hotspot’. Although the insertions are random, there are regions of the chromosome that are more accessible and so these ‘hotspots’ will accumulate a higher incidence of insertion events. One of the *rep* insertions occurs within the promoter and presumably disrupts expression of the gene. The others disrupt the coding sequence and probably result in a truncation of the product, making it non-functional. Despite screening over 9,000 mutants only one other mutation was identified as being synthetically lethal with *recB*. In fact it was not synthetically lethal as previously defined. When tested, the strain gave an unusual phenotype on the indicator plates, yielding large blue and relatively small white colonies (Figure 3.3, row (I), strain SW1030). This ‘synthetically sick’ phenotype suggests that although the double mutant is viable, it is sick when compared to the *recB*<sup>+</sup> strain. Therefore, RecB is not required for survival, but is important to full viability and growth of cells carrying this mutation (Figure 3.3). The mutation also has a phenotype in isolation, causing a significant defect in the growth rate (data not shown) and was identified as an insertion in a gene called *ahpC*.

### **3.2 Mutations in the OxyR regulon are synthetically lethal with *rec* mutations**

The *ahpC* gene is a key component of the OxyR regulon. This regulon deals almost exclusively with the breakdown of H<sub>2</sub>O<sub>2</sub> and although other detoxifying enzymes exist, *ahpC* is believed to encode the primary H<sub>2</sub>O<sub>2</sub> detoxifying enzyme in *E. coli* (Storz *et al.* 1989; Tartaglia *et al.* 1989; Mukhopadhyay and Schellhorn 1997; Ochsner *et al.* 2000; Seaver and Imlay 2001a). The OxyR regulon, which is positively regulated by the OxyR protein in response to oxidative stress, also includes two catalases as well as other detoxifying

proteins that can remove H<sub>2</sub>O<sub>2</sub>. Therefore, in the absence of functional AhpC, the expression of other H<sub>2</sub>O<sub>2</sub> scavenger can be upregulated by OxyR to compensate and reduce oxidative stress (Mukhopadhyay and Schellhorn 1997; Zheng *et al.* 1998; Ochsner *et al.* 2000; Volkert and Landini 2001). However, in the absence of OxyR, the genes in the regulon will either not be expressed at all, or at the very least will not be induced in response to oxidative stress. The complete loss of the oxidative stress response to H<sub>2</sub>O<sub>2</sub>, should be far more detrimental to the cell than the loss of a single H<sub>2</sub>O<sub>2</sub> scavenger.



To test this, a full deletion of the *oxyR* gene was created using the Datsenko and Wanner (2000) recombining method, replacing the *oxyR* gene coding sequence with an apramycin resistance cassette to give strain SW1039 (Table 2.3). The deletion and insertion was confirmed by sequencing, and as expected, the *oxyR* deficient strain showed a marked sensitivity to H<sub>2</sub>O<sub>2</sub> (Kullik *et al.* 1995). The *oxyR* deletion was then introduced as a secondary chromosomal mutation into the *recB* strain carrying a *recB*<sup>+</sup> plasmid, derived from pRC7. When plated onto X-gal, white colonies were extremely rare and very small, suggesting that only cells that retained the plasmid can grow and produce colonies (Figure 3.3, row (I) strain SW1097). The *oxyR* mutation is therefore synthetically lethal with a *recB* mutation. Some interplay between *recB*, which is involved in DNA repair as well as recombination generally, and *oxyR*, which prevents the build-up of H<sub>2</sub>O<sub>2</sub> would seem plausible, as *recB* has been identified as a factor in bacterial resistance to H<sub>2</sub>O<sub>2</sub> (Stohl and Seifert 2006).

H<sub>2</sub>O<sub>2</sub> is probably the principal endogenous ROS and even micromolar concentrations are extremely toxic, causing an array of DNA and other cellular damage (Seaver and Imlay 2001b). In the absence of *oxyR*, H<sub>2</sub>O<sub>2</sub> can accumulate increasing the damage caused. However, the synthetic lethality between *oxyR* and *recB* may be due to a particular activity of *recB*, and not due to the cell's reduced DNA repair ability because of the loss of HR. If recombination is required in an *oxyR* mutant then it would be expected that a *recA* deletion strain, which is almost completely recombination deficient, would also be synthetically lethal with *oxyR*.

To investigate this, a strain carrying a *recA* mutation covered by a pRC7 derived *recA*<sup>+</sup> plasmid (N6121) had secondary *ahpC* and *oxyR* mutations introduced (strains SW1101 and SW1100 respectively). These two mutations had almost identical affects as in the *recB* background when plated on X-gal containing media, *ahpC* being synthetically sick, and *oxyR* synthetically lethal, with *recA* mutations (Figure 3.3, row (II)). This would suggest that the lethality is not specific to a function of RecB, but due to an inability of this strain to utilise recombination. Similar phenotypes were also seen when the *ahpC* and

*oxyR* mutations were placed into strains carrying mutations in *ruvC*, which encodes the only active *E. coli* Holliday junction resolvase (Bennett *et al.* 1993; Mandal *et al.* 1993; Shah *et al.* 1997; Davies and West 1998), or in *recG*, which encodes a dsDNA translocase involved in both replication and recombination (Lloyd 1991; Lloyd and Sharples 1993a; Lloyd and Sharples 1993c; Sharples *et al.* 1999; McGlynn and Lloyd 2001; Gregg *et al.* 2002; Wen *et al.* 2005). Within these genetic backgrounds the *ahpC* mutation only appears to be synthetically sick with *ruvC* (Figure 3.3, row (III), strain SW1145). With *recG* the colonies sector and the blue and white colonies formed are a similar size (Figure 3.4, row (IV), strain SW1124), suggesting that the RecG is not required for survival. However, an *oxyR* mutation is synthetically lethal with both *recG* and *ruvC* on LB (Figure 3.3, strains SW1120 and SW1135). It would appear that in the absence of H<sub>2</sub>O<sub>2</sub> scavengers, cells are forced to rely on HR. Since ROS, in particular H<sub>2</sub>O<sub>2</sub>, cause DNA damage, it would seem likely that recombination is required to repair this DNA damage. Whether HR is needed to repair DSBs, or for the rescue of stalled replication forks is unclear.

If the working hypothesis is true, that synthetic lethality between *oxyR* and the mutations in recombination genes is due to DNA damage caused by H<sub>2</sub>O<sub>2</sub>, then limiting accumulation of H<sub>2</sub>O<sub>2</sub> should alleviate the synthetic lethality and restore a sectoring colony phenotype in the assays used. To reduce the endogenous concentration of H<sub>2</sub>O<sub>2</sub>, two methods were used. Firstly, attempts were made to remove the H<sub>2</sub>O<sub>2</sub> generated, before it can do damage. Although AhpC is the primary endogenous H<sub>2</sub>O<sub>2</sub> scavenger it is not the only one. Catalases are also part of the OxyR regulon and like AhpC they specifically breakdown H<sub>2</sub>O<sub>2</sub> but, unlike AhpC, they only function efficiently when the H<sub>2</sub>O<sub>2</sub> concentration exceeds 1μM (Seaver and Imlay 2001a). By supplementing the agar plates with catalase (150U/ml) it is possible to break down H<sub>2</sub>O<sub>2</sub> present in the media, including any that may diffuse out of a cell across its membrane. Therefore, if the synthetic lethality observed for *oxyR* mutants is due to H<sub>2</sub>O<sub>2</sub> it should be suppressed in the presence of catalase. Indeed, the presence of catalase did restore some viability, allowing white colonies to grow. However, the resulting phenotype suggests that the strains

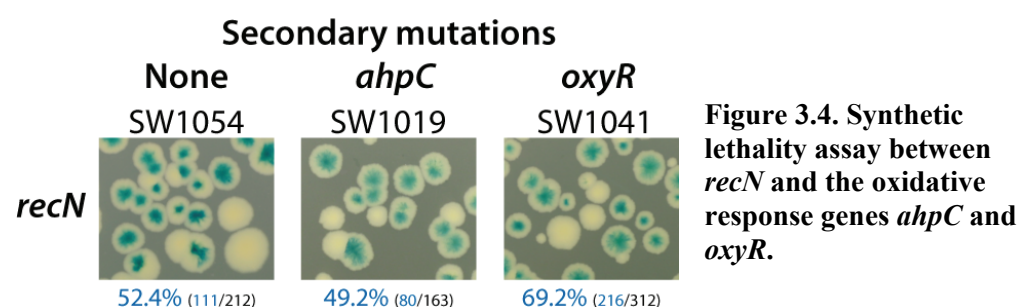
are still synthetically sick, showing the associated phenotype, of large blue and small white colonies, except for the *recG oxyR* mutants which appeared to fully viable with similar sized blue sectoring and white colonies visible (Figure 3.3). The synthetically sick phenotype indicates that while the double mutants can survive, the white *rec<sup>-</sup>* colonies are sick and slower growing compared to those that are *rec<sup>+</sup>*. This suggests that the cells still benefit from having a functional recombination system. The inability to restore full viability may be a consequence of catalase failing to remove all the H<sub>2</sub>O<sub>2</sub> from within the cell. Firstly, catalase can efficiently remove H<sub>2</sub>O<sub>2</sub> only at a relatively high concentration (in excess of 1μM) (Seaver and Imlay 2001b; Seaver and Imlay 2001a). Secondly, the catalase is exogenous, so it can only remove H<sub>2</sub>O<sub>2</sub> that has diffused out of the cells into the media. Although the activity of the catalase will ensure a large concentration gradient across the cell membrane, some H<sub>2</sub>O<sub>2</sub> will remain within the cell. Therefore, it is not possible for exogenous catalase to prevent DNA damage because of H<sub>2</sub>O<sub>2</sub> accumulation, resulting in a requirement for recombination to repair the damage in an OxyR deficient cell.

The second method used to reduce the level of H<sub>2</sub>O<sub>2</sub> within a cell is to limit its production. Endogenous H<sub>2</sub>O<sub>2</sub> is generated as a by-product of electron transfer during respiration, therefore reducing respiration rates will result in reduced production of H<sub>2</sub>O<sub>2</sub> (Beckman and Ames 1998; Storz and Imlay 1999; Seaver and Imlay 2004). Experiments in murine models suggest that calorie restriction can greatly reduce the rate of oxidative phosphorylation, which was linked to reduced rates of ageing (Aruoma *et al.* 1991; Sohal 2002). A similar affect to calorie restriction can be achieved in bacteria by reducing the available nutrients. To do this, cells were grown in LB broth and then plated onto minimal 56/2 salts agar supplemented with Cas-amino acids at 0.1%. Due to the slower growth rate plates were incubated for 72, rather than 48 hours, to allow colony formation prior to analysis. The restricted growth rates seem to largely restore viability to the all the *oxyR* strains, except for *ruvC oxyR* (Figure 3.3). However, the *recB oxyR* and *recG oxyR* double mutants remained a little sick, giving large blue colonies and smaller white colonies. The *recA*

*oxyR* strain had an unusual small, largely white colony phenotype, with only a small, central blue sectorized zone within the colony (Figure 3.3). The small colony size could partly be due to the reduced growth rate of a *recA* deficient strain. Interestingly, the rapid sectoring of the colonies suggests rapid loss of the *recA*<sup>+</sup> plasmid. Therefore, these cells may actually be more viable in the absence of recombination.

The strains were also plated on minimal media, additionally supplemented with catalase. Compared to minimal media without catalase, no affect on the phenotype of the strains was observed, except for the *ruvC oxyR* strain, which was made viable, although it had a synthetically sick phenotype. This suggests that it is catalase which has the effect in this genetic background (data not shown). Growth of the *ahpC* strains on catalase supplemented LB media or on minimal media had no affect with all the strains showing a synthetically sick phenotype, as observed for growth on LB (data not shown).

I used the *recN*<sup>+</sup> derivative of pRC7 (pSW101) to investigate whether RecN was needed to maintain viability in the absence of AhpC or OxyR. The *recN ahpC* and *recN oxyR* cells proved to be viable (Figure 3.4), revealing that RecN is not required in either case. Therefore, it appears that unlike the other *rec* mutants, *recN* is not synthetically lethal with mutations in the oxidative stress response. This is somewhat surprising as ROS are known to induce DNA breakage and RecN is required to repair breaks. Perhaps the rate of break formation is low enough for the cell to cope without RecN.



Mutations are introduced into the chromosome, with the *recN* mutation covered by a pRC7-derived plasmid construct (pSW101), so the strains are effectively *recN*<sup>+</sup>. The effect of introducing the secondary mutations is shown. All colonies were grown on LB agar supplemented with X-gal (6µg/ml) and IPTG (5mM), cells were allowed to grow for 48 hours before being photographed. The total percentage of blue colonies is indicated beneath each panel, with the blue/total colony number shown in parentheses. The strain name is shown above each panel.



### 3.3 Discussion

The synthetic lethal screen initially proved promising for studies into *recN*, especially when its robustness was demonstrated via the isolation of *rep* mutations being synthetically lethal with *recB*. Despite these promising results, the screening did not lead to the discovery of any new *recN* interactions. However, this does not rule out the possibility of as yet unknown interactions, as the screen still has limitations. If the gene is essential, its activity cannot be lost, so it will not be identified. Also, even if an interaction does exist, this does not mean that the absence of the interacting partner would result in lethality in the absence of RecN.

Nonetheless, this study proved fruitful, with the isolation of an *ahpC* mutation that had a synthetically sick phenotype in conjunction with *recB*. This led to the study of the affect of *ahpC* and its regulator *oxyR* when combined with various recombination deficient mutations. This is not the first time that the interplay between DNA repair and oxidative stress regulatory systems has been observed, as discussed previously in context of cellular responses that limit DNA damage (Section 1.2).

All of the recombination genes studied were shown to be synthetically lethal with the *oxyR* deletion and this appears to be due to an inability of these cells to deal with H<sub>2</sub>O<sub>2</sub>, as the addition of catalase to growth media partially restored viability to all the *oxyR rec<sup>-</sup>/ruv* constructs examined. This would also seem to confirm studies that H<sub>2</sub>O<sub>2</sub> is membrane soluble as it must diffuse out of the cell to be degraded by the catalase in the plates (Seaver and Imlay 2001b). However, Seaver and Imlay (2001b) suggest that the permeability of the cell membrane to H<sub>2</sub>O<sub>2</sub> is severely limited, which can lead to substantial differences between intra- and extracellular H<sub>2</sub>O<sub>2</sub> concentrations. Yet, it is clearly adequate to reduce the cellular concentration of H<sub>2</sub>O<sub>2</sub> to a level that is not lethal in the absence of recombination, suggesting that a substantial percentage of H<sub>2</sub>O<sub>2</sub> can diffuse out of the cell. The synthetic lethality between the *oxyR* and *rec* genes highlights the interplay between DNA damage prevention and repair, if one system is inactivated then the cells become dependent on the other.

Intriguingly, it was not possible to restore full viability to all the *oxyR* strains by supplementing the plates with catalase as the strains still showed a synthetically sick phenotype. This suggests DNA damage, and therefore dependence on recombination, still occurred. A possible explanation for this is that in an *oxyR* mutant the *ahpC* gene cannot be expressed. Even when catalase is added not all of the  $H_2O_2$  is removed, this requires the specialised activity of AhpC. Therefore, there is still a low level of  $H_2O_2$ , which leads to chronic DNA damage and in turn, a reliance on DNA repair systems, resulting in the synthetically sick phenotype. The same would occur in the *ahpC rec<sup>-</sup>/ruv* double mutants, explaining their sick phenotype.

The almost identical phenotypes of the *oxyR* mutation in combination with *recA*, *recB* or *ruvC*, and a similar phenotype with *recG*, suggests that the inability to cope with  $H_2O_2$  is due to the cells inability to undergo recombination. HR is crucial to DNA repair, specifically of DSBs and the restart of collapsed replication forks (Kuzminov 1995a; Hoeijmakers 2001; van Gent *et al.* 2001). As  $H_2O_2$  is probably the major ROS, all of which cause DNA damage, it would seem apparent that recombination is needed to repair the DNA damage that results from the increased concentration of  $H_2O_2$  present in the cell. However, ROS will principally cause single- and not double-stranded damage and it is unlikely to generate DSBs directly, which would require recombination to correct. In fact, it is tempting to suggest that the principal issue will be single-stranded lesions leading to replication fork collapse and that recombination will be needed for fork restart. This may explain the less severe phenotype of the *recG* mutants. Generally, they are healthier than the other mutations, with *recG ahpC* being almost normal (although the blue colonies tend to be slightly more common and do not sector as readily) and *recG oxyR* mutants appearing completely healthy in the presence of catalase. RecG is known to promote HR, however it is not essential, but it has also been speculated to rescue stalled replication forks (McGlynn and Lloyd 2000; McGlynn and Lloyd 2001; McGlynn and Lloyd 2002; Donaldson *et al.* 2004). If this is the case it maybe that RecG's non-essential role in HR means that it is only when the concentration of  $H_2O_2$  is

quite high and therefore the DNA damage more severe, as in the *recG oxyR* mutant, that *recG* is essential.

A reliance on HR could also explain the observation that, unlike *oxyR rec<sup>-</sup>* strains, the *oxyR ruvC* strain remains synthetically lethal on minimal media, while *oxyR recA* mutants actually appear to favour a *rec<sup>-</sup>* phenotype. Although H<sub>2</sub>O<sub>2</sub> production is reduced, the overall concentration is still probably higher than when catalase was present in the media. As a result, DNA damage is comparably more common. Alternative repair systems can probably cope with the majority of the single-strand damage caused, and since recombination can be genotoxic, resulting in a variety of genome rearrangements and deleterious mutations, using the alternative repair systems maybe favourable, limiting recombination under these conditions. This would explain the *oxyR recA* strains rapid plasmid loss, giving a completely *rec<sup>-</sup>* strain. However, if recombination is initiated then it must be completed. The activity of *recB* and *recG* is either not essential or can be replaced by other factors and so these strains show only a slight synthetically sick phenotype. However, *ruvC* encodes the sole active cellular resolvase, in its absence crossover products cannot be resolved and these genotoxic recombination intermediates accumulate in the cell. Therefore, it is essential to resolve these intermediates and so the cell requires *ruvC*, resulting in retention of the *ruvC<sup>+</sup>* plasmid and the synthetically lethal phenotype observed on minimal media.

It is worth noting that this is not the first report to associate oxidative stress with a need for repair by HR. Numerous recombination genes, including *recA* and *recBCD* of *E. coli*, have been shown to be important for H<sub>2</sub>O<sub>2</sub> resistance (Linn and Imlay 1987; Asad *et al.* 1997; Konola *et al.* 2000; Bredeche *et al.* 2001; Erill *et al.* 2007). However, in the case of *recA* and *recBCD* mutants the affect is complicated as these mutants are not only compromised for HR, but also SOS induction. It is also noteworthy that OxyR can induce expression of at least one gene known to be involved in DNA repair, *uvrD*, which is involved in NER and MMR (Modrich 1994; Mukhopadhyay and Schellhorn 1997).

What this study does highlight is the reliance cells place on various systems to preserve the genome. DNA damage from ROS is unavoidable and cells therefore have a complex layered defence. Initially trying to prevent ROS accumulation and then repairing any damage that does occur. Even in the absence of one of these systems, a cell remains viable. Oxidative stress is also the primary cause of ageing, a result of cells accumulating DNA damage over time. In fact expression of the bacterial *ahpC* gene in eukaryotic cells can reduce nuclear DNA damage and prevent apoptosis (Lombard *et al.* 2005). Interestingly a recent paper on antibiotic action showed that the major difference between a bacteriostatic and bactericidal antibiotic is that the action of the bactericidal antibiotic results in increased levels of ROS within the cell (Kohanski *et al.* 2007). In an age when antibiotic resistance is becoming ever more prevalent, the potential of any target for antibiotic action is tempting for study. Inhibition of AhpC or OxyR could make cells more sensitive to bactericidal antibiotics and possibly even sensitise cells to antibiotics that have fallen into disuse due to the prevalence of antibiotic resistance. Overall, the importance of ROS scavengers to both the survival of the cell and maintenance of genomic integrity cannot be overstated.

# Chapter 4

## RecN homologs can function in *E. coli*

RecN is an integral part of the SOS response of many bacteria and has a key role in the repair of DSBs. It also appears to be relatively conserved and *recN* mutants have been generated, or identified, in several species, and all show similar phenotypes (Picksley *et al.* 1984a; Van Hoy and Hoch 1990; Sharples and Leach 1995; Funayama *et al.* 1999; Skaar *et al.* 2002; Wang and Maier 2008). Despite this conservation, *recN* is not essential and during normal growth *recN* mutations cause only minor phenotypic affects. It is only when a cell is forced to rely on HR, for instance to repair DSBs, that the absence of *recN* becomes apparent (Picksley *et al.* 1985b; Sargentini and Smith 1986; Simic *et al.* 1991; Meddows *et al.* 2005; Erill *et al.* 2007). This chapter will explore how well RecN is conserved across species and what implications this may have for our understanding of its role in the cell.

### 4.1 RecN homology

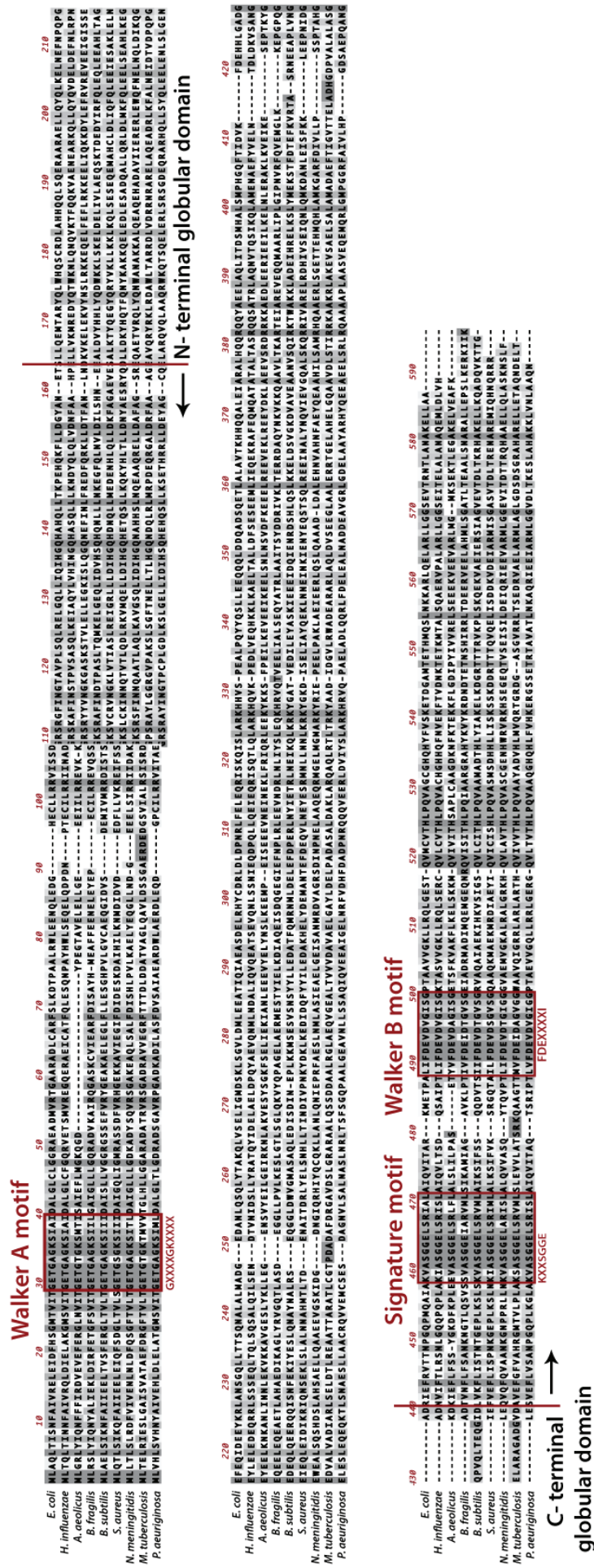
A search of the TIGR database (<http://cmr.tigr.org>) shows that there is an annotated RecN homolog in all of the currently sequenced bacterial genomes (404 are currently complete as of the 2<sup>nd</sup> December 2008), with the exception of the *Thermatoga* species. The absence of RecN in this group could be due to incomplete annotation of these genomes. However, BLAST searches of the *Thermatoga* genomes and proteome, against several RecN sequences, including *E. coli* and *A. aeolicus*, as well as a consensus RecN sequence that was generated, failed to find a RecN homolog. The searches did identify a *recF*-like gene and the *Thermatoga* SMC protein. Both are SMC-like proteins and so share homology with RecN (Section 1.5). RecN is therefore not just well conserved, but essentially ubiquitous in bacterial genomes, suggesting that its function is also highly conserved and must be of great importance to a cell.

Alignment of the RecN protein sequences from several species shows that the proteins share considerable homology. However, this appears to be limited to the N- and C-terminal domains of RecN (Figure 4.1). These correspond to the globular domains that form the head of an SMC protein. The two globular domains of RecN contain well conserved motifs that have been identified as Walker A, Walker B and Signature motifs (Meddows 2002). These motifs can form between them a functional ATP binding and hydrolysis pocket, if brought together (Walker *et al.* 1982). The residues that are essential for ATPase activity are particularly well conserved, including K35 in the Walker A box, essential for ATP hydrolysis, and the FDE residues of the Walker B box, required for ATP binding.

The Walker motifs are located at opposite termini of the protein (Figure 4.1). However, to function they must be in proximity. The eukaryotic SMC proteins have a similar arrangement, with the head domains separated by a long coiled-coil region, which can fold back on itself, due to a central flexible 'hinge' domain, thus bringing the globular heads together (Haering *et al.* 2002; Hirano 2005b). The central region of the RecN proteins show limited sequence conservation. This could represent the coiled-coil region, the sequence of which is of little consequence, as long as the structural form is conserved. There is, therefore, limited amino acid conservation in this region. It could even be advantageous for an organism to modify the coiled-coil depending on the environment it inhabits (Trivedi *et al.* 2006). Unfortunately, no conserved hinge domain could be identified. However, by examining the alignment data it was possible to estimate where the interfaces between the conserved globular and coiled-coil domains are located (for *E. coli* RecN, residues 1-152 are predicted to be the N-terminal domain and 390-553 the C-terminal domain) (Figure 4.1).

Since sequence homology appears to be restricted to the globular domains, and the probable location of these can be predicted, it is therefore possible to align the globular domains separately, which is useful as RecN is small for an SMC family member. When the head domains of *E. coli* RecN and several known SMC proteins were aligned, homology was shown to be

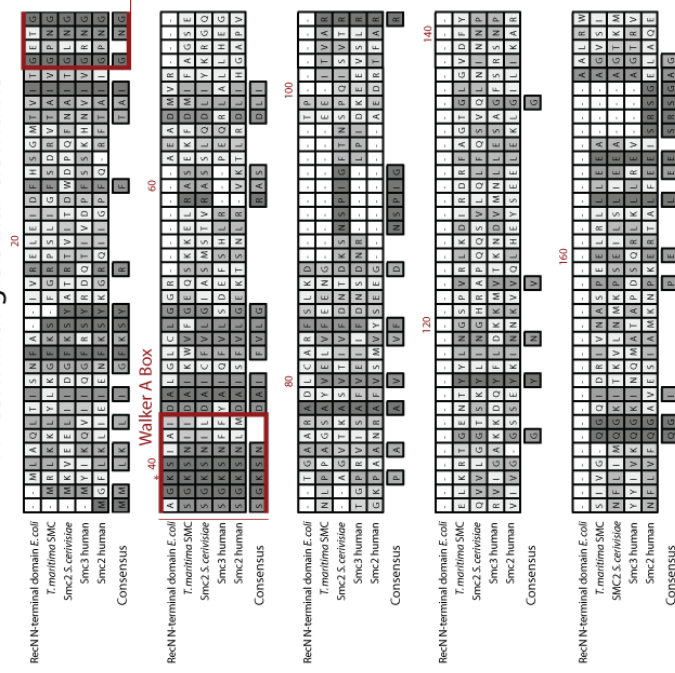
extremely limited, almost exclusively in fact to the Walker and signature motifs (Figure 4.2). However, even SMC proteins from the same species show limited homology. It would seem that the best way to characterise an SMC protein is via structural and not sequence homology, as has been done for RecN (Sharples and Leach 1995).



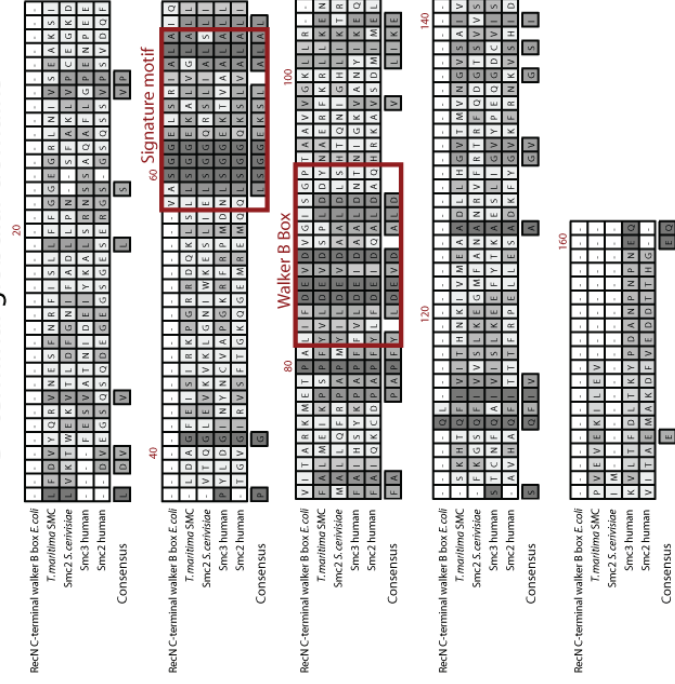
**Figure 4.1. Alignment of amino acid sequences of RecN homologs, using ClustalW 1.4.**



## N-terminal globular domains



## C-terminal globular domains



**Figure 4.2. Alignment of amino acid sequences of the globular domains of *E. coli* RecN and known SMC proteins, using ClustalW1.4.**

Motifs important to ATPase are boxed in red. The intensity of shading indicates the conservation of a residue (darkest being the most highly conserved). Conservation is largely limited to the essential motifs.

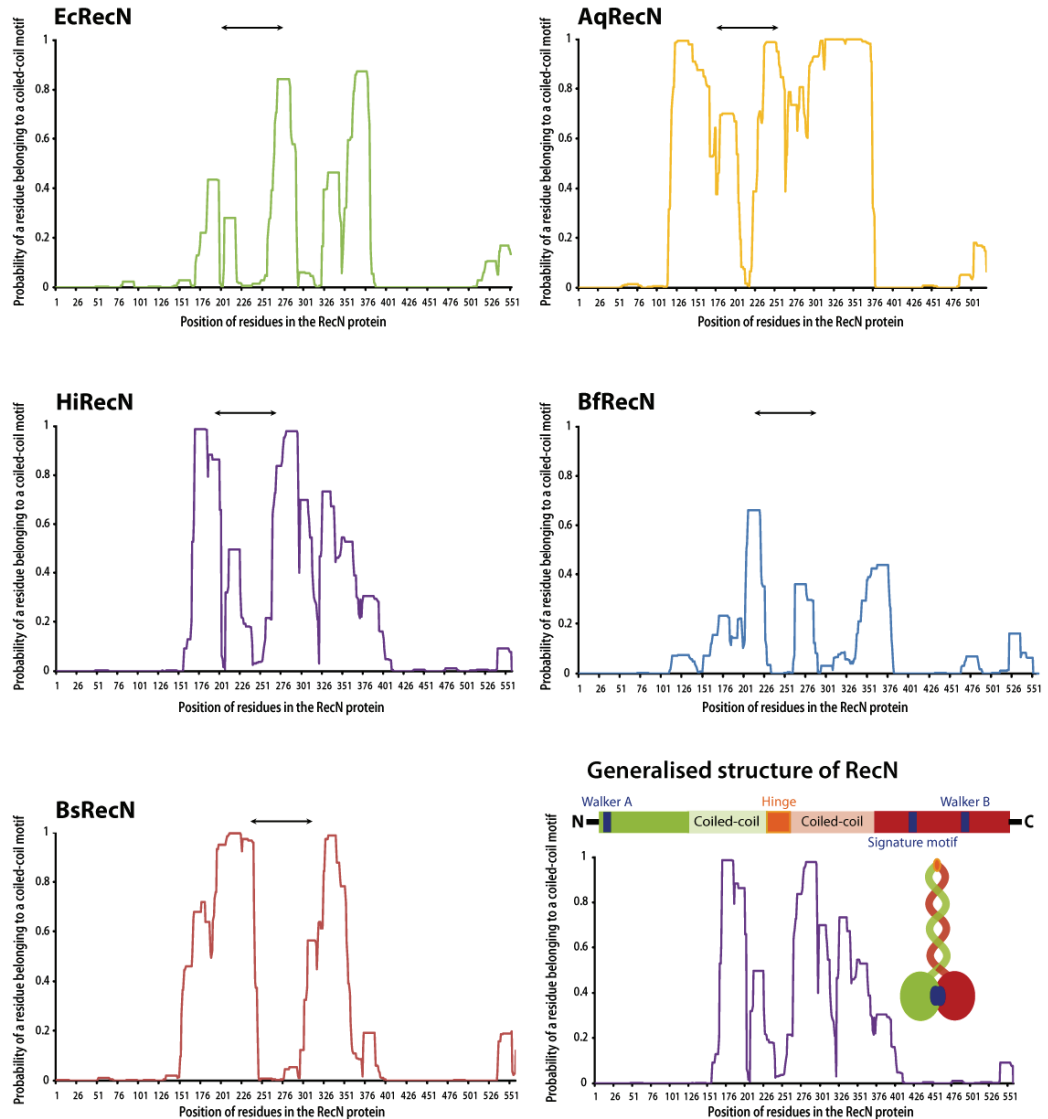
## 4.2 RecN protein structure

SMC proteins have a distinctive structure. Each SMC protein possesses two globular domains, one at the N-terminus and one at the C-terminus. These contain the conserved ATPase motifs and are separated by a long anti-parallel coiled-coil motif with a central flexible hinge domain. The hinge domain allows the protein to fold back on itself and it is via this intra-molecular folding that the globular domains are brought into contact to create the 'head domain'. It is believed that the folded SMC proteins usually dimerise, and the association of the two head domains causes functional ATP binding pockets to form (Melby *et al.* 1998; Haering *et al.* 2002; Hirano and Hirano 2002; Hopfner and Tainer 2003; Cobbe and Heck 2004; Nasmyth and Haering 2005). RecN proteins share sequence homology (Figure 4.1). However, this is limited to the globular, head forming, domains. In contrast the homology between RecN and other SMC proteins is remote, limited almost entirely to the ATPase motifs (Chapter 4). It is therefore suggested that it is structural and not sequence homology that is important in recognising SMC family members and that perhaps this structural homology also represents functional homology (Sharples and Leach 1995).

One of the most recognisable features of an SMC protein is its long coiled-coil region, the presence of which can be predicted via BLAST searches as was used to assign residues to the globular domains (figure 4.1). It is also possible to predict the likelihood of an individual residue belonging to a coiled-coil motif *in silico*. Using the COILS software (Lupas 1991), it is possible to predict the locations and extent of the coiled-coil regions more accurately (Figure 4.3). It is worth noting that the alpha helices that comprise a coiled-coil are themselves extremely common structural elements of proteins and so false positives can occur (Lupas *et al.* 1991; Jones 1999; Offer *et al.* 2002). From the predictions of the COILS software the likelihood of residues belonging to a coiled-coil was plotted (Figure 4.3). As expected, the data suggest that the N- and C-terminus of the protein, previously assigned to the globular domains, have a very low probability of being part of a coiled-coil, while the central region has a high probability. In three of the plots (HiRecN,

AqRecN and BsRecN) there appears to be two distinct regions of coiled-coil motif, separated from each other by a region with a very low probability of being a coiled-coil (Figure 6.11). As the globular domains of RecN appear to be of a similar size to those of known SMC proteins (Figure 4.2), it would appear that the ‘missing’ residues of RecN would be part of the coiled-coil region. This means that if RecN does function as a cohesin-like protein and forms a ring structure to trap DNA, this ring, assuming RecN forms a dimer, would have a much smaller diameter (Haering *et al.* 2002; Gruber *et al.* 2003). As bacterial chromosomes are generally smaller and have less higher order structure than their eukaryotic counterparts, it is feasible that two or more DNA strands could fit within a RecN ring, allowing RecN to function as a cohesin-like molecule (Chapter 1). However, evidence from *B. subtilis* suggests RecN forms an octameric, rather than a dimeric, protein complex (Kidane *et al.* 2004).

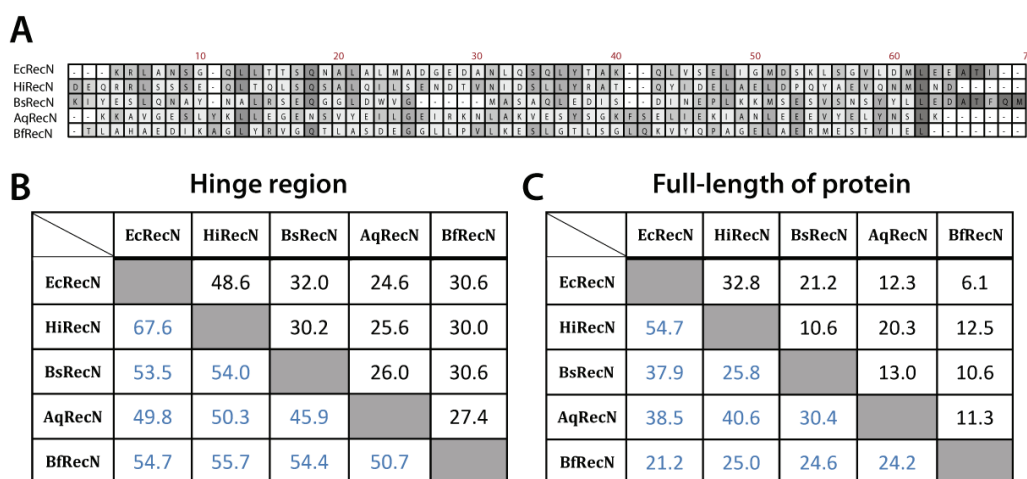
The identification of a central region of low coiled-coil probability gives a possible location for the, as yet unidentified, hinge domain of RecN (Figure 4.3). The hinge is crucial to the function of SMC proteins as it is essential to folding of the molecule and believed to mediate several activities including dimerisation and loading onto DNA (Hirano *et al.* 2001; Hirano and Hirano 2002; Gruber *et al.* 2006; Hirano and Hirano 2006). This would suggest that some sequence conservation in this region should be expected. However, the hinge domain of EcRecN and BfRecN remains indistinct, with several central regions of low coiled-coil probability. By isolating and aligning the region independently I tried to detect the hinge domain and any conserved sequences and identify the location of EcRecN and BfRecN hinges. Therefore, I performed sequence alignments of the 70 residues in and around the proposed hinge domains (Figure 4.4).



**Figure 4.3. Graphs showing the likelihood of each individual residue belonging to a coiled-coil motif for several RecN proteins.**

The likelihood of a residue belonging to a coiled-coil motif was predicted using the COILS software (Lupas 1991) and plotted against its position within RecN. This is shown for all the RecN proteins which have been part of this study and also BsRecN. The black arrow shows the 70 amino acid sequence aligned as part of the hinge studies and identified as the likely hinge region. The generalised structure of RecN compares the coiled-coil prediction for HiRecN to the proposed structure of RecN and how it can fold bringing the globular domains together as well as what regions of the plot belong to which domain of the protein.

The alignment of the sequences around the proposed hinge regions, allowed the identification of the likely EcRecN and BfRecN hinges (Figure 6.11, black arrows). When these sequences were aligned they can be seen to share a reasonable degree of similarity, with the hinge showing an elevated level of sequence similarity and identity compared to the full-length protein (Figure 6.12 B and C). When the full-length proteins were aligned (Figure 4.1) no hinge domain was detectable. It may be possible in the future to use this information to target site directed mutagenesis on residues likely to be essential to hinge function, especially as *in vivo* and *in vitro* assays are now available for RecN function.



**Figure 4.4. Relationship of RecN proteins using clustalW v1.83.**

- A)** The sequence alignment for the 70 residues that correspond to the hinge regions of the RecN proteins, as proposed by the coiled-coil predictions (Figure 6.11). The alignment was performed in MacVector 9.5.2 using clustalW v1.83.
- B)** Table showing the percentage of sequence identity (black) and similarity (blue) of the 70 residue of the hinge region alone.
- C)** Table showing the percentage of sequence identity (black) and similarity (blue) for the full length of the RecN proteins.

The homology between the globular domains of RecN proteins suggests conservation and it is plausible that this is adequate to allow them to share functionality. It was therefore decided to test RecN homologs for their ability

to function in *E. coli*, in place of the *E. coli* RecN (EcRecN). It was also hoped that these homologs could then be used for *in vitro* studies, which has not been feasible for the *E. coli* protein.

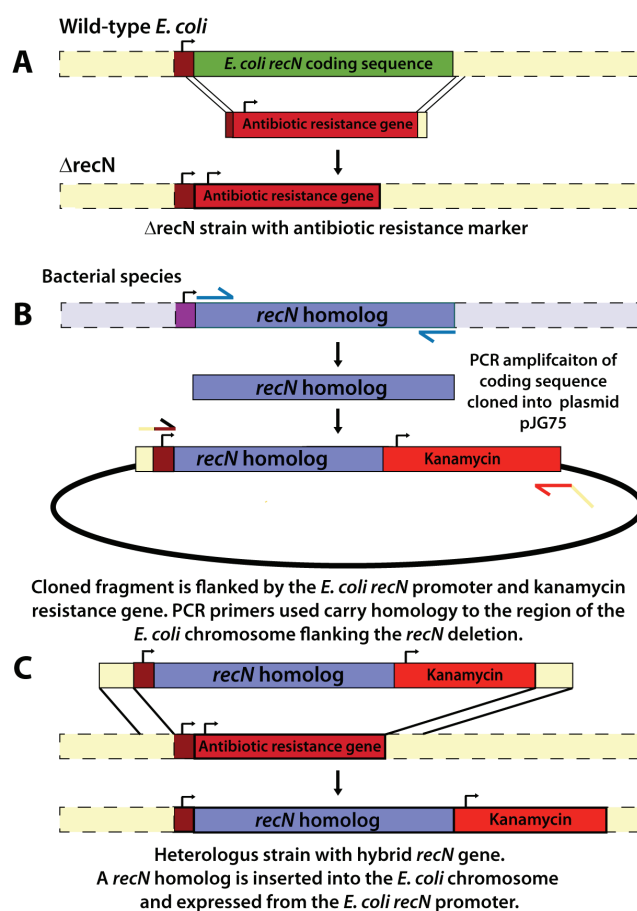
### 4.3 Heterologous genetics

Initially, three bacterial RecN homologs were tested for their ability to replace EcRecN *in vivo*. These homologs were cloned from *Haemophilus influenzae* (HiRecN), a bacterium closely related to *E. coli*, *Bacteriodes fragilis* (BfRecN), an obligate anaerobe that lives in the gut and *Aquifex aeolicus* (AqRecN), a hyperthermophile, found in hot water springs, growing at temperatures in excess of 90°C. *Aquifex aeolicus* is part of the earliest branching families within the bacterial domain (Deckert *et al.* 1998).

To test complementation, strains were constructed (Dr. Jane Grove, University of Nottingham) whereby the coding sequence of the EcRecN was replaced with that of the homolog on the chromosome, whilst leaving the *E. coli* *recN* promoter sequences intact. This produced a heterologous strain that expresses a RecN homolog in place of EcRecN and eliminated any issues associated with expression from a plasmid or from inducible promoter systems. Replacement of EcRecN first required a full deletion of the coding sequence, using the recombineering method described by Datsenko and Wanner (2000) to insert a trimethoprim resistance marker in its place (Datsenko and Wanner 2000). Once the deletion was created, the same protocol was used to insert the homologous *recN* coding sequences, replacing the trimethoprim resistance cassette with the alternative RecN protein coding sequence linked to a kanamycin resistance marker downstream.

To do this, PCR primers were designed to allow cloning of the coding sequences of the RecN homologs, into a plasmid (pJG75). This allowed the cloned sequence to be flanked by the *E. coli* *recN* promoter sequence at the 5' end and a kanamycin resistance cassette at the 3' end. The resulting plasmid provided a template for generating PCR products for recombineering (Datsenko and Wanner 2000) the hybrid gene into the chromosome of *E. coli* (Figure 4.5). The linked kanamycin resistance gene allows movement of the

hybrid *recN* construct into different *E. coli* genetic backgrounds via P1 transduction (Chapter 2). All of the chromosomal integrants were sequenced to confirm their identity and correct integration. This showed a silent, C to T transition, in codon 39 (valine) of the *H. influenzae recN* sequence, when compared to that published in GenBank. These heterologous strains are designed to express a RecN homolog in place of the *E. coli* protein, with expression regulated by the *E. coli recN* promoter sequence (Figure 4.5, strains made are listed in Table 2.3, plasmid constructs in Table 2.5).



**Figure 4.5. Constructing heterologous RecN *E. coli* strains.**

Primer positions are shown, with colours suggesting homology. Dashed outlines indicating chromosomal DNA.

**A)** The *E. coli recN* coding sequence (green) is deleted and replaced with an antibiotic resistance cassette (red). Using the Datsenko and Wanner method (Chapter 2). To create a  $\Delta recN$  strain this is then used in the construction of heterologous *recN* strains

**B)** The *recN* coding sequence from another bacteria (purple) is amplified by PCR and cloned into vector pJG75, between the *E. coli recN* promoter and a kanamycin resistance gene. This plasmid is used to generate a PCR product with the *recN* homolog and the

kanamycin resistance gene cloned onto a DNA fragment flanked by sequences homologous to the *E. coli* chromosome and containing a kanamycin resistance gene as a marker

**C)** The PCR product is used to replace the chromosomal antibiotic resistance cassette. This creates a heterologous strain that expresses the AqRecN (JIG652), BfRecN (JIG655) or HiRecN (JIG645) (regulated by the EcRecN promoter (dark red). As a control an *E. coli recN* construct was produced with the kanamycin resistance downstream and shown to behave as a wild-type strain (JIG725).

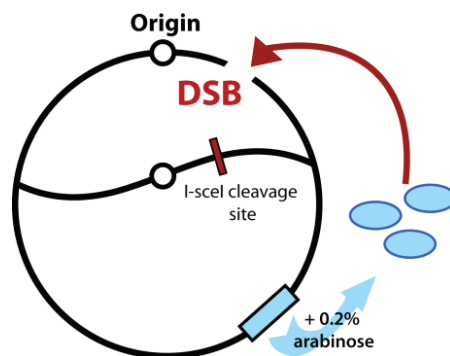
#### 4.4. *H. influenzae* RecN can function instead of *E. coli* RecN *in vivo*

The major phenotype of a RecN deficient cell results from an inability to repair DSBs, which manifests as sensitivity to any agent that causes DSBs. Therefore, to ascertain if the RecN homologs could complement for EcRecN *in vivo*, the heterologous strains were assessed as to how well they survive DSBs generated by I-SceI.

##### 4.4.1 The I-SceI cleavage system

This system allows DSBs to be generated in an inducible, site-specific manner, using I-SceI homing endonuclease from yeast, which cuts DNA at a specific 14bp site, not found naturally in the *E. coli* chromosome (Meddows *et al.* 2004). Both cleavage sites and an arabinose inducible I-SceI expression cassette have been engineered into the *E. coli* chromosome at specific loci (Figure 4.4). Addition of arabinose to these strains during growth allows expression of I-SceI (Guzman *et al.* 1995), causing a DSB at the cleavage site (Meddows 2002).

The effect of transient induction of I-SceI on cell survival is readily measured by comparing the viable cell count per ml of a glucose fed culture with a sample of the same culture fed with arabinose. Studies have shown that even an essentially wild-type *E. coli* strain is sensitive to DSBs generated by I-SceI, with only 10-15% of cells surviving under the assay conditions established. A  $\Delta recN$  strain is yet more sensitive, with cell survival below 0.1% (Grove *et al.* 2008).



**Figure 4.6. I-SceI induced cleavage of the *E. coli* chromosome.**

Addition of arabinose induces expression of I-SceI (blue circles) from an expression cassette (blue square). Cleavage occurs at defined sites (red rectangle). Pre-replicated regions provide a template for repair (Meddows 2002).

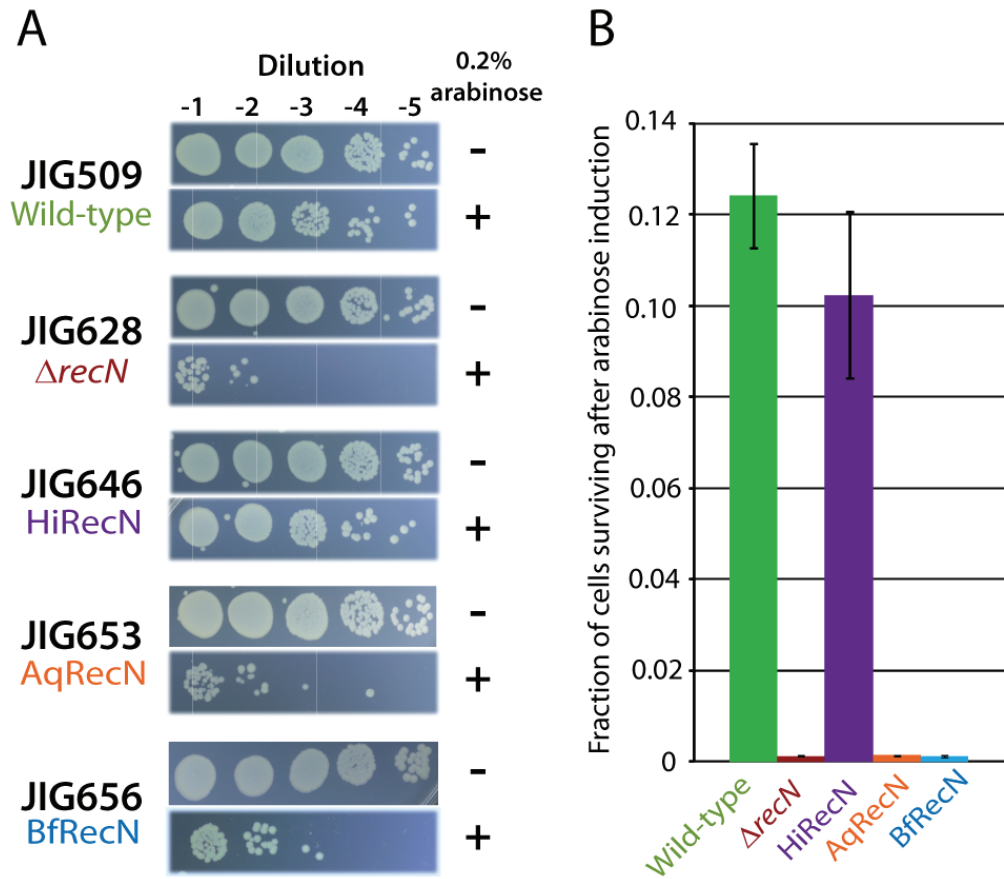


#### 4.4.2 Survival of heterologous strains after transient I-SceI expression

To assess the ability of the hybrid *recN* genes to function in *E. coli*, the genes were introduced into a strain carrying an I-SceI expression cassette in the *attB* gene locus and I-SceI cleavage site engineered into *lacZ*. This allowed me to assess whether the heterologous RecNs could promote repair of DSBs.

All the strains are sensitive to DSBs generated by I-SceI (Figure 4.7). However, the survival of a  $\Delta recN$  strain (JIG628) is reduced over a 100-fold compared to the wild-type (JIG509) (Figure 4.7). When the heterologous strains were assayed it is readily apparent that only the strain carrying the *H. influenzae recN* hybrid gene (JIG646) has a viability comparable to wild-type (Figure 4.7 A and B).

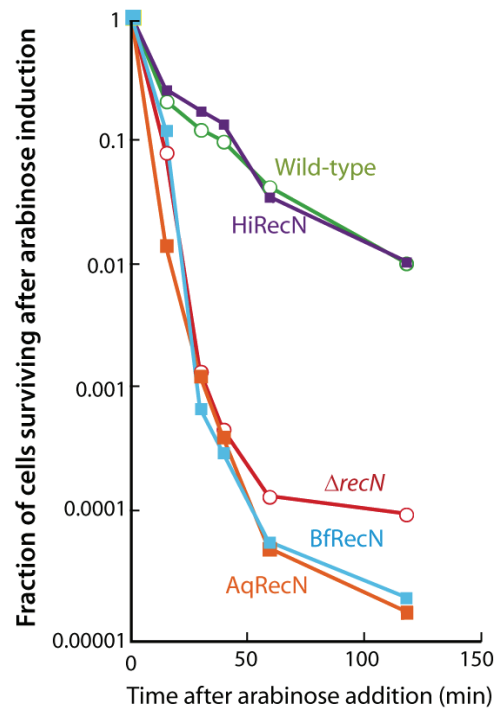
These results were further analysed by experiments where the I-SceI exposure was continuous over a two hour period and the decrease in viable cells measured (Figure 4.8). In the absence of RecN the viable cell count diminished rapidly compared to the wild-type *E. coli*. However, in this case it can be clearly seen that the strain carrying the *H. influenzae recN* hybrid gene survived as well as the wild-type. Meanwhile the other hybrids showed severely compromised survival and were clearly more sensitive to the DSBs caused by I-SceI. It would appear that only the RecN from *H. influenzae* can complement for its *E. coli* homolog and facilitate repair of the DSBs induced by I-SceI.



**Figure 4.7. I-SceI viabilities of *E. coli* strains expressing RecN homologs.**

All strains carry an I-SceI expression cassette and a cleavage site. The strains used are otherwise wild-type strain (JIG509, green n = 9),  $\Delta recN$  (JIG628, red, n = 4), or heterologous strains expressing HiRecN (JIG646, purple n = 4), AqRecN (JIG653, orange n = 4) or BfRecN (JIG656, light blue n = 3).

- A)** Serial dilution and spotting of 10 $\mu$ l of culture grown with or without transient induction of I-SceI for 35min. Error bars are standard error of the mean for combined data set number of repeats is shown for each strain by the n value.
- B)** Graph showing the average viabilities after transient I-SceI exposure.



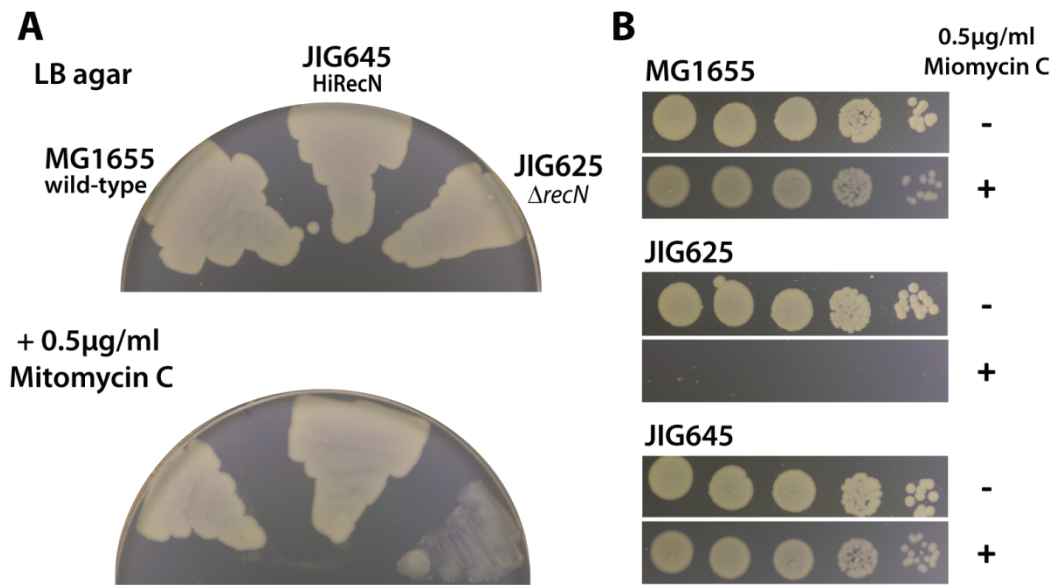
**Figure 4.8. Long term I-SceI viability of heterologous strains over two hours.**

All strains carry an I-SceI expression cassette and a cleavage site. The strains used are otherwise wild-type strain (JIG509, green n = 8),  $\Delta recN$  (JIG628, red n = 4), or heterologous strains expressing HiRecN (JIG646, purple n = 3), AqRecN (JIG653, orange n = 3) or BfRecN (JIG656, light blue n = 3). Error bars are not shown as they are too small on the logarithmic scale used

#### 4.4.3 Complementation of mitomycin C sensitivity

It has been noted that *recN* mutations show variable sensitivity to different DSB causing agents, when compared to recombination deficient strains such as *recA* mutants, which are consistently extremely sick after exposure to such agents. This may reflect different structures or types of break (Picksley *et al.* 1984a; Kosa *et al.* 2004). Therefore, although HiRecN could complement for EcRecN in respect to I-SceI induced breaks the question remains can it complement when DSBs are formed by other methods? The sensitivity of *recN* mutants to mitomycin C is well documented not only for *E. coli*, but several other species as well (Iyer and Szybalski 1963; Picksley *et al.* 1984a; Mascarenhas *et al.* 2006; Wang and Maier 2008). Although mitomycin C does not cause DSBs directly, it is useful to ascertain the functionality of the RecN homologs strains *in vivo*.

The hybrid *recN* genes were transduced into an MG1655 *E. coli* background and these constructs grown to an OD<sub>650</sub> of 0.48 before testing for sensitivity to mitomycin C. A simple streak test showed that the HiRecN protein could confer resistance (Figure 4.9 A). A semi-quantitative spot dilution revealed that this protein is as effective as the wild-type *E. coli* RecN, conferring resistance to mitomycin C (Figure 4.9 B).

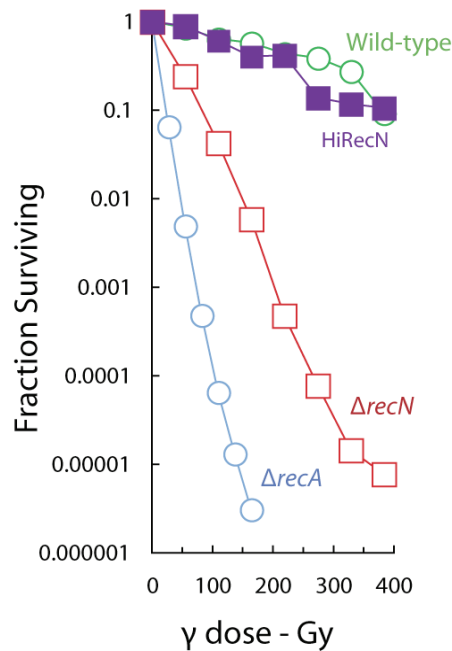


**Figure 4.9. Sensitivity of *E. coli* strains to the DNA damaging agent mitomycin C.**

- A)** 10µl of culture at OD<sub>650</sub> of 0.48 were streaked directly onto LB agar plates with and without mitomycin C at a concentration of 0.5µg/ml. Wild-type (MG1655), *H. influenzae* hybrid (JIG645) and  $\Delta recN$  (JIG625) strains were tested.
- B)** Serial dilution  $10^{-1}$  -  $10^{-5}$  of culture at OD<sub>650</sub> 0.48 was spotted (10µl) onto LB  $\pm$  mitomycin C at 0.5µg/ml.

#### 4.4.4 Complementation of Ionising Radiation sensitivity

One of the first phenotypes identified for a *recN* mutation was sensitivity to IR (Picksley *et al.* 1984a), which causes DSBs in a dose dependent manner (Ward 1975; Ward 1988). *E. coli* strains carrying *recN* mutations have an intermediate sensitivity to IR, when compared to a wild-type and a  $\Delta recA$  strain. When the heterologous HiRecN *E. coli* strain was exposed to IR, it was as resistant as a wild-type *E. coli*. It appears that HiRecN functions as effectively as EcRecN, fully complementing an array of phenotypes associated with RecN deficiency (Figure 4.10).



**Figure 4.10. Sensitivity of the hybrid *H. influenzae* RecN and other *E. coli* strains to IR**

Comparison of MG1655 (wild-type, open green circles), the heterologous HiRecN expressing strain (closed, purple squares, JIG646), an *E. coli*  $\Delta recN$  (open red squares, TRM160) and a  $\Delta recA$  strain (open light blue circles, N4279). Survival is plotted against increasing doses of IR. All experiments were repeated three times, no error bars are visible on the logarithmic scale used.

## 4.5 Discussion

### 4.5.1 Conservation of RecN

Sequence analysis suggests that RecN is well conserved within the bacterial kingdom. Although, actual sequence conservation is limited to the globular domains, with the ATPase motifs being particularly well conserved. There is considerable structural homology between both the RecN proteins and SMC proteins. The preservation of RecN throughout bacterial species, suggests it has an important role in the cell. However, mutants and deletions can be generated, which under normal growth conditions are healthy, except for a moderate induction of the SOS response (Simic *et al.* 1991). RecN shares limited sequence homology, but does have structural homology, with the SMC family of proteins (Sharples and Leach 1995). This is despite this protein being much smaller (approximately a third the size) of many of the SMC proteins. However, even closely related SMC proteins have limited sequence homology, with their relationships being determined by structural and functional homology. The only way to determine if RecNs structural resemblance to an

SMC protein relates to a functional homology is to identify its activities *in vitro* and *in vivo* and compare them to those of SMC proteins.

#### 4.5.2 Conservation of RecN function between homologs

Out of the RecN homologs studied, only HiRecN could replace the function of EcRecN. The inability of the *A. aeolicus* and *B. fragilis* RecN to function in *E. coli* could be due to several factors. Both of these proteins show greater sequence divergence from the *E. coli* protein, than HiRecN. This could suggest that RecN has conserved interaction partners in the cell and the more divergent RecN proteins simply cannot interact with the *E. coli* RecN partners. This would fit with the data on SMC proteins, which form complexes with non-SMC proteins (Section 1.5) that are essential for function (Haering and Nasmyth 2003; Hirano and Hirano 2004; Nasmyth and Haering 2005).

An alternative hypothesis could be that *B. fragilis* and *A. aeolicus* RecN are inadequately expressed, at least in an active form, within *E. coli*. Expression of RecN is massively induced during the SOS response, making it one of the most abundant proteins in *E. coli* under these conditions (Finch *et al.* 1985a). Even though the *E. coli* promoter regulates the expression of the RecN homologs and consequently the mRNA should be expressed as if they were the *E. coli* protein, this does not mean that the subsequent translation and correct protein folding occur. *B. fragilis* shares similar codon usage to *E. coli*. Therefore, *E. coli* cells should be able to translate *B. fragilis* RecN mRNA efficiently. However, issues with expression could be particularly relevant for the AqRecN. *A. aeolicus* is a hyperthermophile, found in hot water springs. It grows in conditions where temperatures exceed 90°C and makes use of several codons that are rare in *E. coli*. This could severely limit the expression of *A. aeolicus* proteins within *E. coli* cells. Secondly, its proteins have to be active and stable at elevated temperatures and it could simply be that although the AqRecN is produced, it cannot function at 37°C. Testing activity *in vivo* at elevated temperatures (42°C) showed no improvement (data not shown). However, even 42°C is still relatively low for *A. aeolicus* growth.

### 4.5.3 Summary

RecN is an almost ubiquitous within the largest domain of life on earth. Yet how RecN functions and what role it plays in DSB repair remain unknown. The study of RecN should provide a useful new insight into the repair of DSBs and genome dynamics in bacteria. To date analysis of RecN's function in *E. coli* has been limited to genetics as the protein cannot be isolated. However, here I have identified a RecN homolog from *H. influenzae* that appears to fully substitute for *E. coli* RecN. It could therefore be studied instead of the *E. coli* protein. Any *in vitro* activities discovered should be applicable to the *E. coli* protein (Chapters 5 and 6).

# Chapter 5

## Purification of RecN

Despite the discovery of the *recN* gene in *E. coli* over 20 years ago and the subsequent studies into its genetics, remarkably little is known about the biochemical activities of RecN. The only major biochemical data comes from studies of BsRecN, and to date these have not been fully conclusive or reconciled with the data available on the *E. coli* protein (Section 1.5). However, the purification of functional *E. coli* RecN has not been successful, largely due to its insoluble nature (Meddows 2002; Neher *et al.* 2006).

Since BsRecN has been purified successfully (Kidane *et al.* 2004; Sanchez and Alonso 2005), it seemed possible that the insolubility of EcRecN is specific to this protein and not an inherent property of RecN proteins in general. This raised the possibility of purifying and studying other EcRecN homologs. In Chapter 4, the ubiquity of RecN in bacteria was highlighted, along with the ability of the *recN* gene from *H. influenzae*, to complement *in vivo* for a deletion of the *E. coli* gene. This means that studies of the *H. influenzae* protein (HiRecN) could provide data that is applicable to EcRecN.

Even though HiRecN was the only homolog to function in place of EcRecN, purification of all three RecN homologs described in Chapter 4 was attempted. These three have different physical properties (Table 5.1), therefore it was hoped that one or more might be purified, and in a functional form. If more than one RecN could be purified, their comparison would allow definition of intrinsic, and species-specific attributes. In this chapter the successful expression and purification of all three RecN homologs; HiRecN, *A. Aeolicus* RecN (AqRecN) and *B. fragilis* RecN (BfRecN) is described, with particular reference to the purification of HiRecN.



**Table 5.1. The physical properties of *E. coli* RecN homologs.**

RecN <sup>a</sup> homolog	Nucleotide number	Amino acids	Mass (kDa)	Predicted pI	% amino acid identity to <i>E. coli</i> RecN
<i>E. coli</i>	1662	553	61.37	5.24	100
<i>H. influenzae</i>	1677	558	62.63	4.59	48
<i>A. aeolicus</i>	1563	520	60.43	4.74	20
<i>B. fragilis</i>	1686	561	62.93	4.84	34
<i>B. subtilis</i>	1731	576	64.48	4.80	30

<sup>a</sup> These proteins have either been purified, or their purification attempted. Those in red were examined in this study.

## 5.1 Expression of RecN proteins in *E. coli*

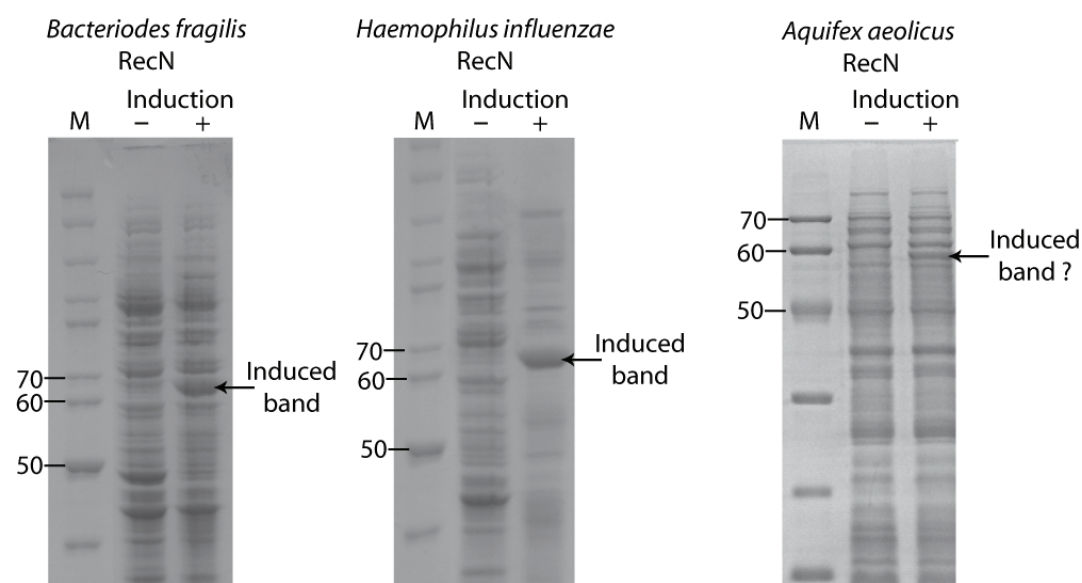
The proteins were expressed in *E. coli* using a variety of plasmid constructs (Table 5.2), production of which is described in Chapter 2. All of the constructs were sequenced to confirm their identity. Initially, the RecN homologs were screened to determine if they could be expressed in *E. coli*. Cultures were grown at 37°C, in the presence of appropriate antibiotics and trials were performed by growing two 8ml cultures in parallel. When they reached an OD<sub>650</sub> of 0.4, one of the cultures was induced (Table 5.2) to express the RecN homolog before incubating for a further four hours. The cells were then harvested by centrifugation and lysed by re-suspending the cell pellets in 1ml of SDS-PAGE running buffer. Samples of both induced and non-induced cultures were then analysed by SDS-PAGE. Both BfRecN and HiRecN were induced (Figure 5.1), with protein bands of about the correct size being visible on the gel. Unfortunately, although the AqRecN appeared to be expressed, its yield was relatively low compared to the other two RecN proteins (Figure 5.1).

**Table 5.2 Expression vectors used to express RecN homologs in this study.**

Expressed protein	Expression vector	Vector backbone	Mode of induction	Expressed in strain
<i>H. influenzae</i>	pJG71	pLau17	Arabinose <sup>a</sup>	MG1655
<i>A. aeolicus</i>	pTRM129	pET3a	IPTG <sup>b</sup>	BL-21(DE3) pLysS
<i>B. fragilis</i>	pJG99	pLau17	Arabinose <sup>a</sup>	MG1655

<sup>a</sup> 0.2% final concentration

<sup>b</sup> 5mM final concentration



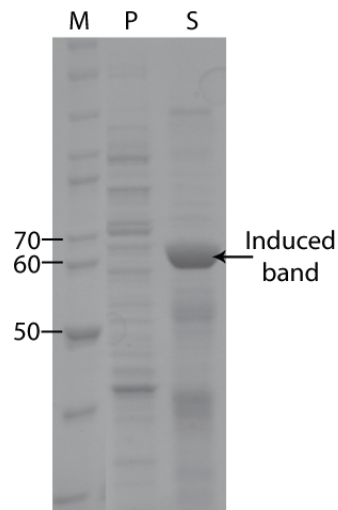
**Figure 5.1. Small-scale overexpression of RecN homologs.**

Total proteome of cultures uninduced (-), or induced (+) to express the RecN protein indicated. Size markers are in lane M with relevant sizes marked (kDa).

## 5.2 Purification of RecN from *H. influenzae*

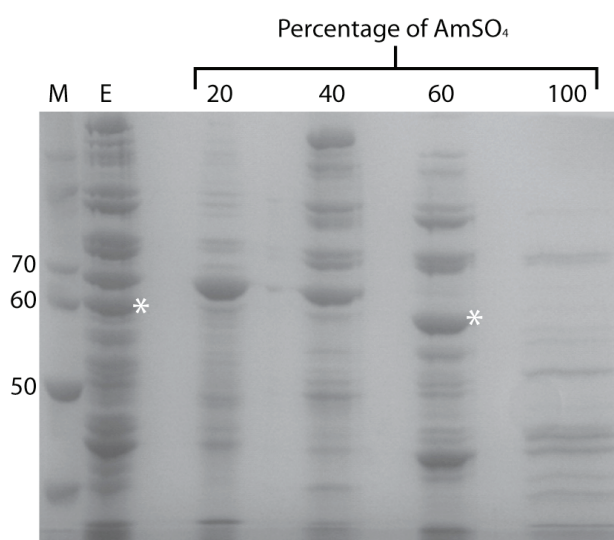
To determine whether HiRecN is soluble, induced cells were harvested, lysed by sonication in lysis buffer (50mM Tris-HCl, pH7.5, 150mM NaCl) and then clarified by high-speed centrifugation (15,000g, 30min). When samples of both the supernatant and the resuspended cell pellet were run on a polyacrylamide gel, the HiRecN was found almost exclusively in the soluble fraction (Figure 5.2).

The solubility of a protein in solution can be decreased by the addition of ammonium sulphate ((NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>). Different proteins will precipitate at different (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> concentrations, allowing a cell extract to be fractionated (Chapter 2). The amount of (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> required to precipitate HiRecN was determined in a small-scale trial. The soluble protein fraction, from cells induced to express HiRecN, was isolated. (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> was added to reach a set percentage of saturation, and proteins allowed to precipitate for 1 hour at 4°C; the solution was then clarified by centrifugation before increasing the (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> concentration and repeating precipitation. Samples of the pellet after each concentration were then analysed by SDS-PAGE. The HiRecN was found in the 60% saturation pellet, but is not in the 40% pellet (Figure 5.3).



**Figure 5.2. The soluble and insoluble proteome of *E. coli* cells induced to express HiRecN.**

Polyacrylamide gel showing insoluble proteins, present in the pellet (Lane P) and soluble proteins are present in the supernatant (Lane S). Size markers are in Lane M, with appropriate size markers shown (kDa).



**Figure 5.3. Ammonium sulphate precipitation of *E. coli* cell extracts containing HiRecN.**

Cell extract (lane E) was treated to raise the  $(\text{NH}_4)_2\text{SO}_4$  concentration to the percentage of total saturation indicated. Samples of precipitated proteins analysed on a SDS-PAGE gel. HiRecN protein (\*) is present in the cell extract and 60% sample. Size markers are in lane M, with relevant sizes marked (kDa).

To scale up the purification, cells from 400ml of culture, grown aerobically at 37°C and induced to express HiRecN for four hours after reaching an  $\text{OD}_{650}$  of 0.4, were harvested and the pellet stored overnight at -80°C. The cells were then resuspended in 0.01 cell volumes of lysis buffer and lysed by sonication on ice, after which the soluble fraction was isolated by centrifugation (15,000g, 30min). This was precipitated with  $(\text{NH}_4)_2\text{SO}_4$  at a concentration of 40%. The pellet discarded and then the remaining cell extract precipitated with  $(\text{NH}_4)_2\text{SO}_4$  at a concentration of 60% saturation, this time the pellet was retained. The  $(\text{NH}_4)_2\text{SO}_4$  precipitation was performed at 4°C, with each precipitation being left for one hour before isolation of the pellet.

HiRecN was then further purified by chromatography at 4°C, using an ÄKTA FPLC system (GE Healthcare). All of the columns used were purchased pre-packed from GE Healthcare, unless stated otherwise. The  $(\text{NH}_4)_2\text{SO}_4$  pellet was resuspended in buffer A (50mM Tris-HCl pH7.5, 1mM EDTA, 1mM DTT) supplemented with 1 M  $(\text{NH}_4)_2\text{SO}_4$  and loaded onto a 20ml Hiprep 16/10 FF Phenyl-sepharose hydrophobic interaction column. Proteins were eluted, with a gradient of  $(\text{NH}_4)_2\text{SO}_4$  from one to zero Molar, across ten column volumes (CV), in buffer A. HiRecN eluted between 500-350mM  $(\text{NH}_4)_2\text{SO}_4$  (conductivity of 82 – 62mS/cm) (Figure 5.4).

Fractions containing HiRecN (Figure 5.4) were pooled, and diluted with buffer A so as to reduce the salt concentration, allowing HiRecN to bind

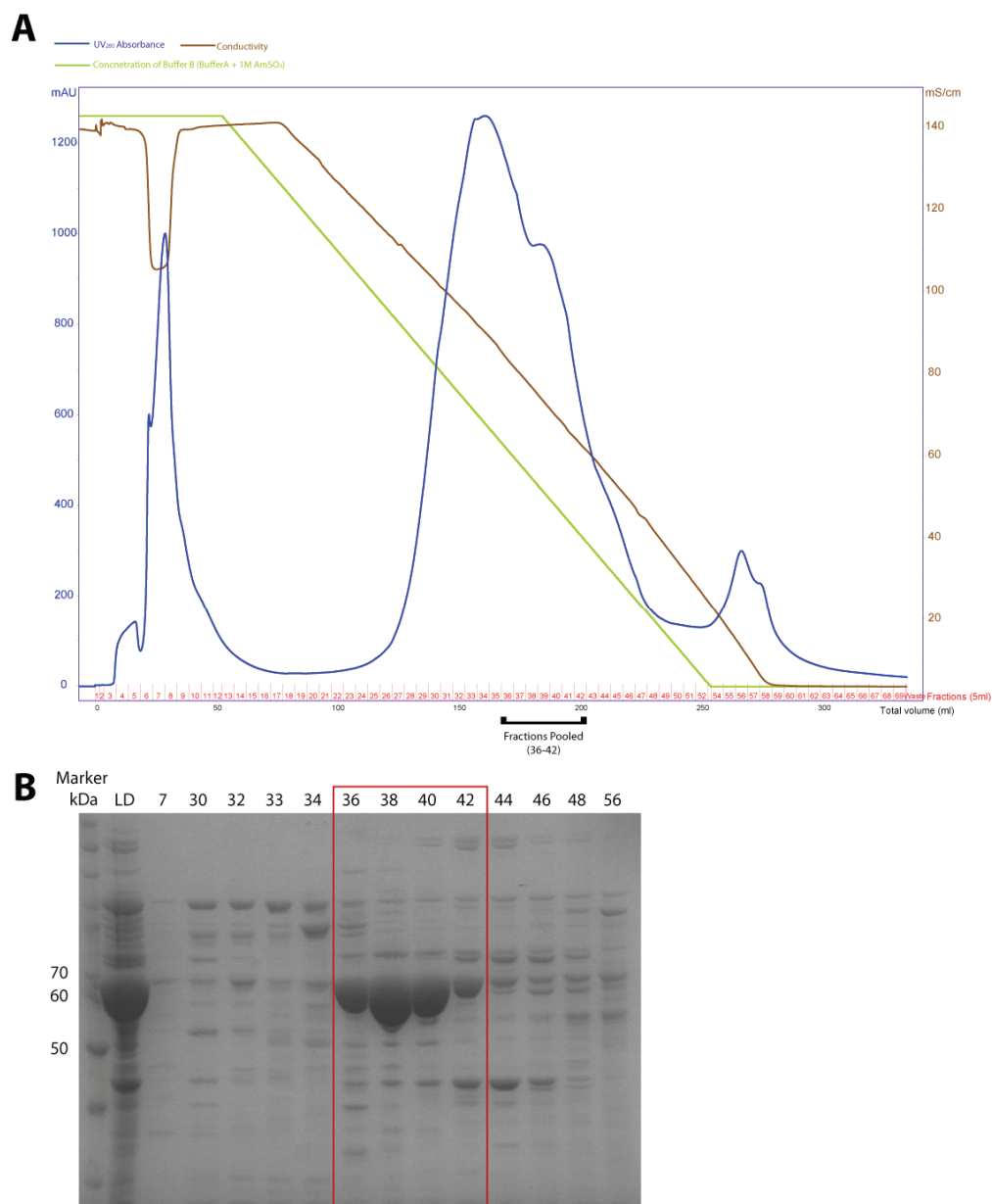
2x5ml Hitrap Q HP ion exchange columns. Typically, this required dilution so that the final conductivity was less than 20mS/cm. The diluted fractions were then loaded onto the column, the column washed with 0.2M NaCl in buffer A, and proteins eluted with a salt gradient, over 10 CV, of 0.2-1.0M NaCl in buffer A. There were two distinct peaks of material eluted from the column, as indicated by the UV absorbance trace (Figure 5.5). The first occurred between 350-380mM NaCl (conductivity of 38-45mS/cm) and the second at 450-500mM NaCl (58-62mS/cm conductivity). SDS-PAGE analysis indicated that only the first peak contained HiRecN. The second contained no protein material (by SDS-PAGE). Ethidium bromide staining of an agarose gel revealed that these fractions contained DNA (data not shown) and it seems therefore that the observed UV absorbance was due to the elution of DNA (Figure 5.5).

The fractions containing HiRecN (Figure 5.5) were pooled once again and diluted with buffer A, so that the conductivity was <10mS/cm. This allowed purification of the HiRecN using a 5ml heparin column. HiRecN will not adhere to the heparin column, but several contaminants do (Figure 5.6). However, after the protein is washed through the sample required concentrating, so as to reduce total volume for handling and loading onto the next column. Therefore, a 1ml Hitrap HP Q column, was placed in series with the heparin column. The HiRecN binds to the Q column after it has passed through the heparin column. The heparin column was then removed and the proteins eluted from the 1ml Q column. HiRecN was found in the first peak of the UV trace at 350-380mM NaCl (Conductivity of 58-62mS/cm) (Figure 5.6).

The pooled fractions containing HiRecN (Figure 5.6) were then run on a HiPrep 16/60 Sephacryl S-300 HR size exclusion column (GE Healthcare), in gel filtration buffer (50mM Tris-HCl, pH7.5, 100mM NaCl). This column separates proteins and protein complexes based on their apparent size; the larger they are, the sooner they elute. It is possible to use this to estimate a protein's size, from its elution volume, by comparing to the elution volumes of known size standards. HiRecN elutes early, in the first peak seen (Figure 5.7), at a volume of 40.0-54.1ml (median of peak is at 46.3ml). By comparison to

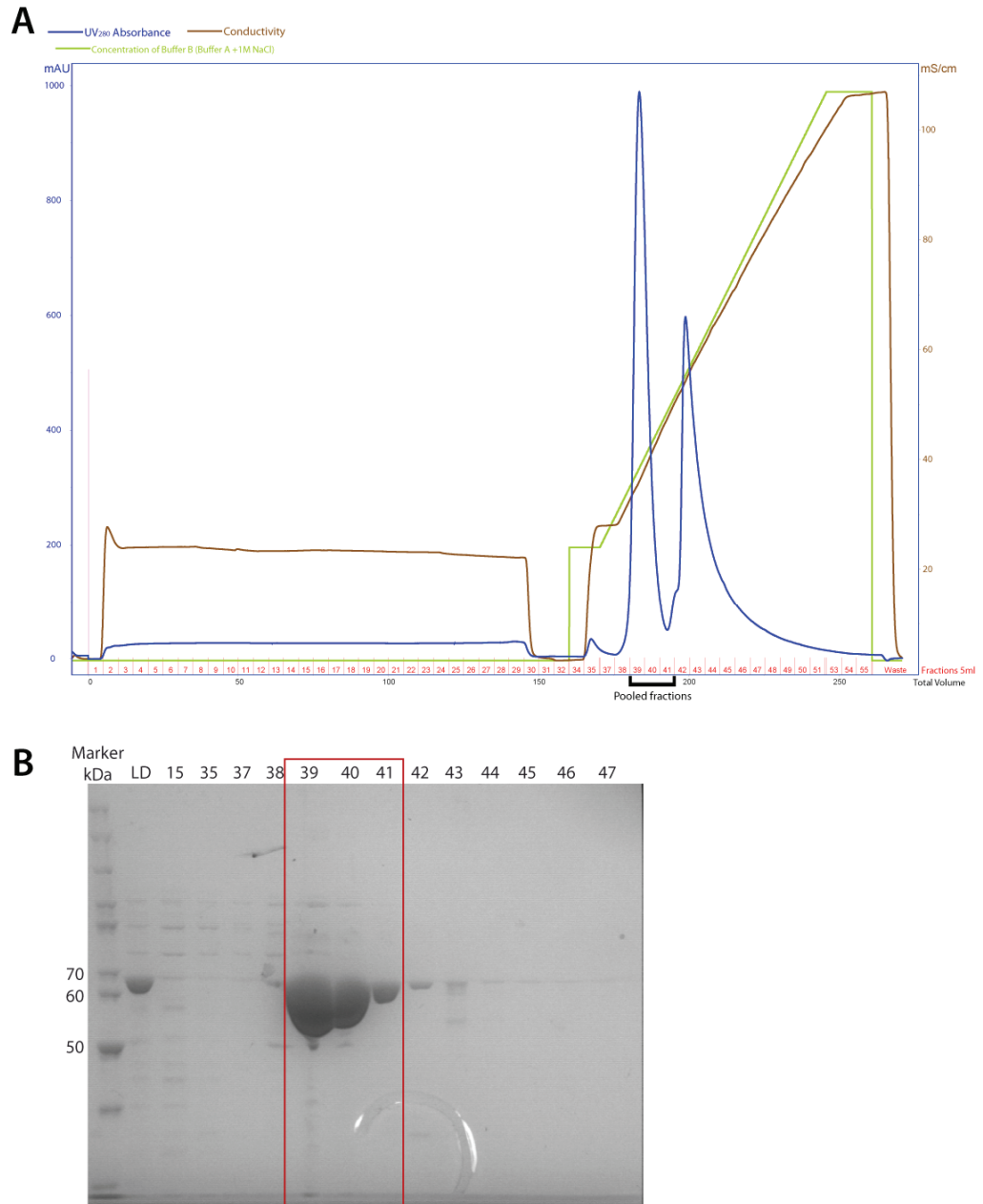
the elution volume of known proteins (Figure 5.7) it is possible to estimate the size of these complexes and the elution volume of HiRecN suggests a protein complex of 650-156kDa (325kDa) in size. The formation of large RecN complexes would fit with data on BsRecN, which forms large protein complexes *in vivo* and *in vitro* (Kidane *et al.* 2004; Sanchez *et al.* 2007a). Unfortunately, gel filtration does not provide a definitive measure of protein size as it assumes the protein is globular. However, SMC proteins have elongated forms with the coiled-coil domains protruding away from the molecule (Melby *et al.* 1998; Hirano 2002). This would increase their apparent size, or Stokes radius, drastically. The increased Stokes radius of RecN will therefore make it appear to run on a gel filtration column as a far larger protein than it actually is. Yet, gel filtration does suggest that RecN forms some type of Homeric complex in solution. Importantly, the size exclusion column did further the purification of HiRecN, removing two contaminating proteins, both of which ran as smaller species on the column, but one actually appears larger upon SDS-PAGE analysis (Figure 5.7).

As a final purification and concentration step, fractions containing HiRecN were loaded onto a Mono Q HR 5/5 ion exchange column. The column was washed with 0.2M NaCl in buffer A and bound proteins eluted across a gradient of 0.2 – 0.5M NaCl, over 20CVs in buffer A. This time there was only a single UV absorbance peak, which was shown to contain the HiRecN, which appears to be pure on an SDS-PAGE gel (Figure 5.8). The fractions containing pure HiRecN were pooled and dialysed, against a storage buffer (50mM Tris-HCl, pH7.5, 100mM NaCl), in which the protein is stable for at least two weeks, at 4°C (data not shown). Glycerol could not be used as a cryoprotectant for freezing, as it was found to interfere with activity (Chapter 6). Once dialysed, the protein concentration as determined by the Bio-Rad protein assay (Bio-Rad) with BSA as the standard was approximately 4mg/ml. The identity of the protein was confirmed by MALDI-PMF fingerprinting (University of Nottingham, BSAU). This compares the trypsin digest profile of the protein, as seen by mass spectrometry, to those in the SWISS-Prot and TrEMBL databases.



**Figure 5.4. Purification of HiRecN using a 20ml Hiprep 16/10 FF Phenyl-sepharose column.**

- A)** Trace showing the UV absorbance (blue line), conductivity (brown line) and the concentration of buffer B (buffer A + 1M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>) (green line). The UV absorbance peaks suggest where proteins are eluted.
- B)** Samples from fractions coinciding with the UV peaks were loaded onto an SDS-PAGE gel, as indicated. Size markers, with relevant sizes marked (kDa), are shown. LD indicates load material and the fractions pooled for further purification (containing the HiRecN) are boxed in red on the gel and indicated on the trace (A).

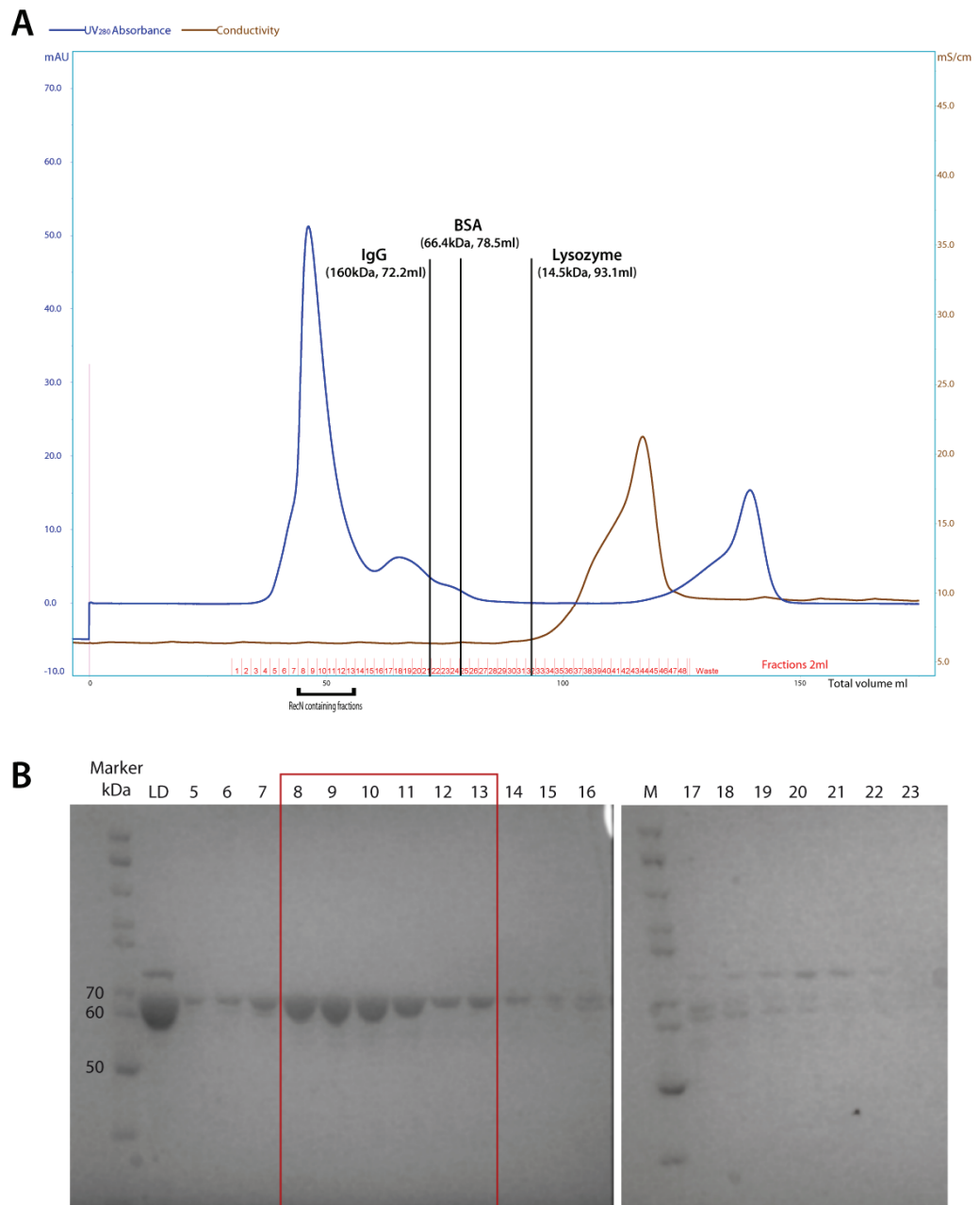


**Figure 5.5. Purification of HiRecN using 2 x 5ml Q HP columns.**

- A)** Trace showing the UV absorbance (blue line), conductivity (brown line) and the concentration of buffer B (buffer A + 1M NaCl) (green line). The UV absorbance peaks suggest where proteins are eluted.
- B)** Samples from fractions coinciding with the UV peaks were loaded onto an SDS-PAGE gel, as indicated. Size markers, with relevant sizes marked (kDa), are shown. LD indicates load material and the fractions pooled for further purification (containing the HiRecN) are boxed in red on the gel and indicated on the trace (A).

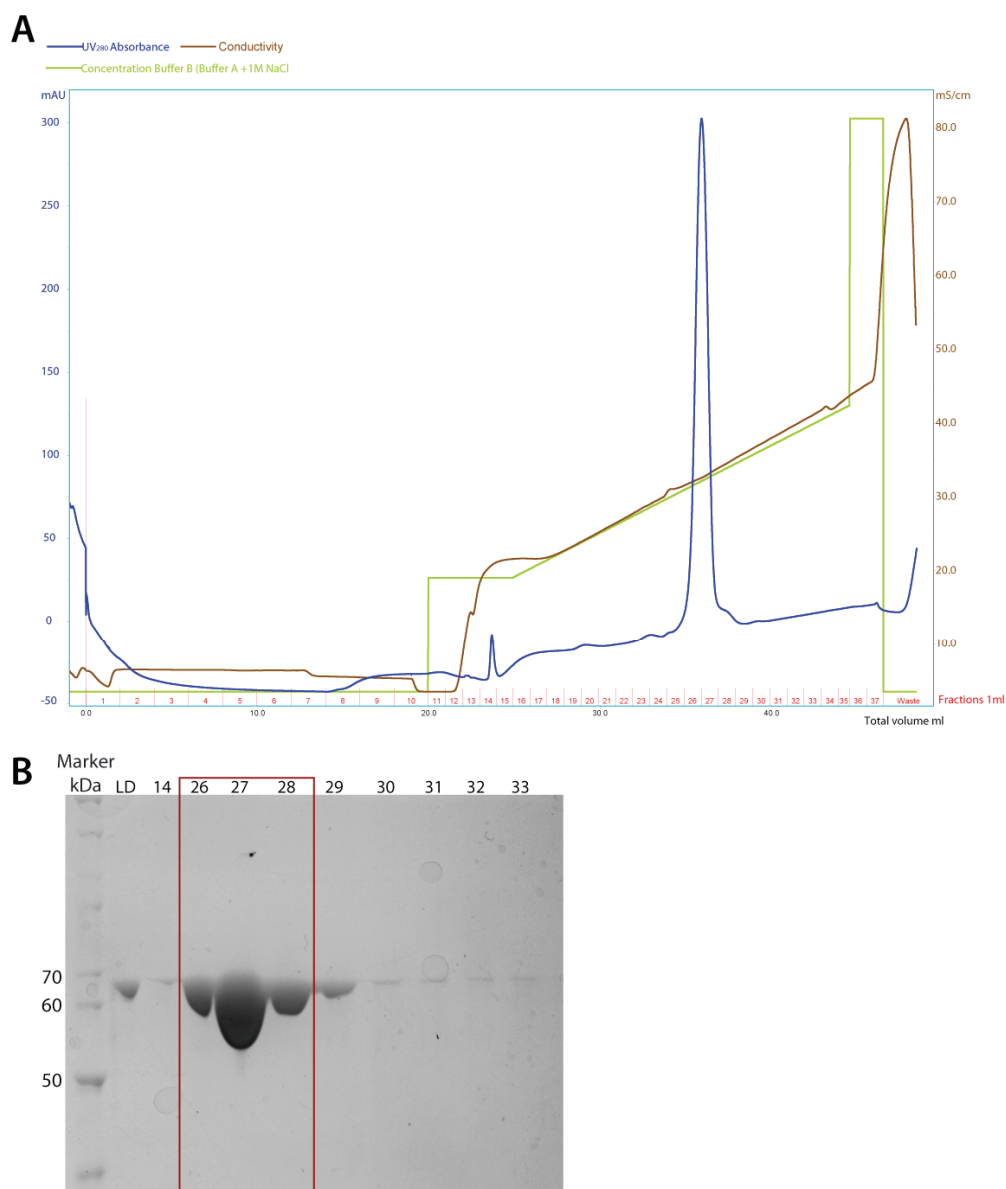






**Figure 5.7. Purification and determination of apparent size of HiRecN using a 120ml Hiprep 16/60 Sephacryl S-300 HR column.**

- A)** The UV trace peaks as proteins are eluted (blue line); the conductivity (brown line) is also shown. The centre of elution peaks of known protein standards used to calibrate the column are shown as vertical black lines along with details of the standards used.
- B)** Samples from fractions coinciding with the UV peaks were loaded onto an SDS-PAGE gel, as indicated on the gel. Size markers with relevant sizes marked are in lane 1. LD indicates the load material. The fractions, which were pooled (containing HiRecN) are boxed in red on the gel and indicated on the trace (A).



**Figure 5.8. Final stage of HiRecN purification, using a 1ml Mono Q HR 5/5-anion exchange column.**

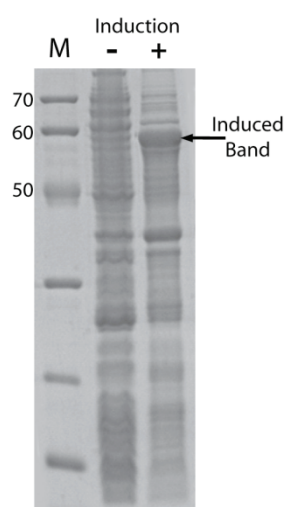
- A)** The UV trace (blue line) suggests were proteins are eluted, while conductivity (brown line) and the concentration of buffer B (buffer A +1M NaCl) (green line), are also shown.
- B)** Samples from fractions coinciding with the UV peak were loaded onto an SDS-PAGE gel as indicated. Fractions pooled and kept are boxed in red. Lane 1 contains size markers with relevant sizes marked (kDa). LD indicates the load material.

## 5.3 Purification of the other RecN homologs

Both BfRecN and AqRecN were successfully expressed within *E. coli* cells, albeit inefficiently in the case of AqRecN (Figure 5.1).

### 5.3.1 *Aquifex aeolicus* RecN purification

The limited expression of AqRecN could be due to its coding sequence containing several codons that are rare in *E. coli*. Expression can be improved by increasing the final cell density and rate of growth (Derewenda 2004). To achieve better cell growth, and hopefully improve expression of AqRecN, the cells were grown in a fermenter (Electrolab, Fermac 360/310). This allowed regulation of growth conditions; such as temperature, as well as improving aeration and agitation of the media, which was a rich (2xYT) broth. All these factors can improve a cell's growth (Neubauer *et al.* 1992). Induction was performed in the same manner as the small-scale trials, yet in this case expression of AqRecN was clearly induced after four hours (Figure 5.9).

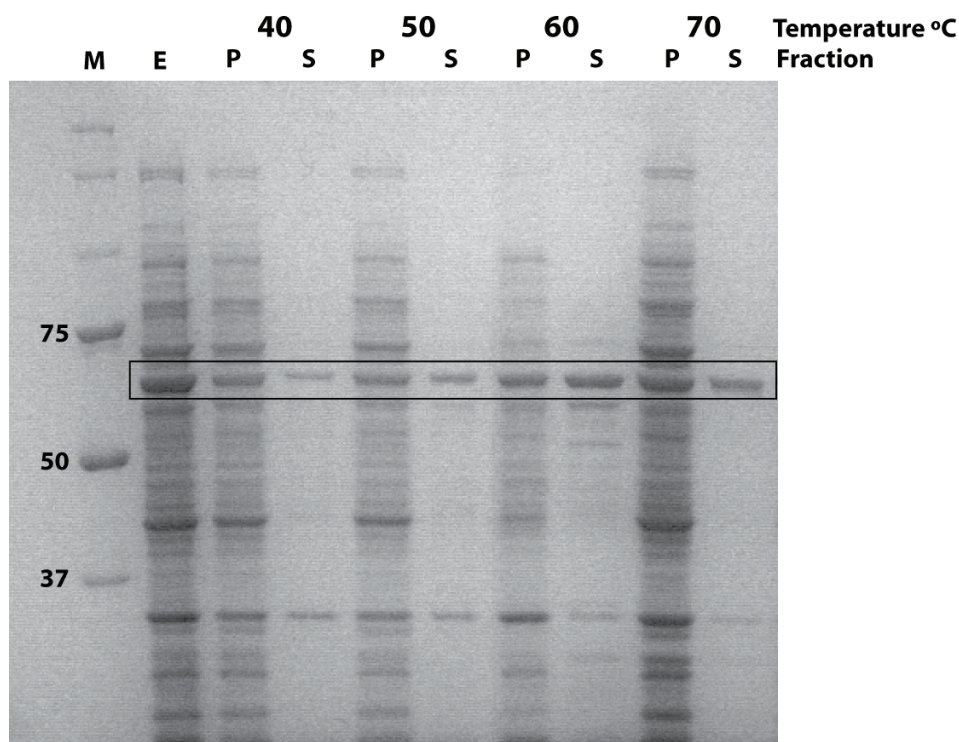


**Figure 5.9. Expression of *A. aeolicus* RecN, in *E. coli* grown in a fermenter.**

Cells were grown and induced in 2xYT media, at 37°C. Size markers are in lane M (kDa), uninduced culture (-) and induced (+) were run with the induced band (AqRecN) being indicated.

Upon cell lysis, the AqRecN was shown to be largely insoluble (data not shown), and thus impossible to purify in a functional form directly. However, *A. aeolicus* proteins should be thermostable. Heating of the cell extract was therefore trialled as a method to liberate AqRecN into the soluble fraction. Heating for twenty minutes prior to clarification was shown to increase the proportion of AqRecN in the soluble fraction and also denatures *E.*

*coli* proteins, making them insoluble and thus acting as a purification step. Increasing the temperature generally increased the quantity of AqRecN liberated, and the percentage of *E. coli* proteins denatured, with the highest yield of AqRecN occurring at approximately 60°C (Figure 5.10).



**Figure 5.10. Effect of heating on the soluble proteome of *E. coli* cells overexpressing AqRecN.**

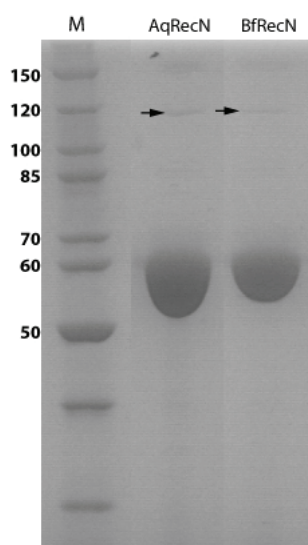
Total cell extract (E) was heated to the temperatures specified, for 20min, and then clarified (15,000g, 30min) with the cell pellet (P) and supernatant (S) then being run on the SDS-PAGE gel. AqRecN is boxed, with size markers (lane M) shown and appropriate sizes (kDa) indicated.

To purify AqRecN, 2.5 litres of induced, fermenter-grown cells were resuspended in lysis buffer, lysed by sonication and then heated to 60°C for twenty minutes, before being clarified by centrifugation. The supernatant was then precipitated with  $(\text{NH}_4)_2\text{SO}_4$ , the 40% pellet discarded and the 60% pellet resuspended and purified on an ÄKTA FPLC as described for HiRecN. Elution from the columns was comparable to that observed for HiRecN, except from the 20ml Hiprep 16/10 FF Phenyl-sepharose column, where AqRecN eluted at  $(\text{NH}_4)_2\text{SO}_4$  concentrations of less than 100mM, in buffer A (conductivity was

<40mS/cm. The protein was purified to a gel pure state (Figure 5.11) and could be concentrated in excess of 7mg/ml, as determined by the Bio-Rad protein assay with BSA as the standard.

### 5.3.2 *Bacteriodes fragilis* RecN purification

BfRecN was also successfully purified. When overexpressed it was, like HiRecN, found almost exclusively in the soluble fraction. However, it did elute from some of the columns differently. Like AqRecN, it eluted from a 20ml Hiprep 16/10 FF Phenyl-sepharose column at  $(\text{NH}_4)_2\text{SO}_4$  concentrations of less than 100mM, and on gel filtration eluted much earlier than the other two RecN proteins, suggesting it forms a much larger protein complex (Table 5.3). The reason for this, and its significance, is currently unclear. The protein appears to be pure as seen on a SDS-PAGE gel (Figure 5.11) and has a good yield (Table 5.3), although the maximum protein concentration is somewhat lower than AqRecN, being approximately 4-5mg/ml as determined by the Bio-Rad protein assay.



**Figure 5.11. Purified AqRecN and BfRecN analysed on a SDS-PAGE gel.**

No contaminants are visible, despite the high load of protein except for the faint band, shown by black arrows, which was sequenced and shown to be RecN protein. This may represent a dimeric species of the proteins. Size markers are in lane M with relevant sizes indicated. In both cases, the purified and concentrated protein was run on the gel, with equal load volume of each. The identity of both proteins was confirmed by MALDI-PMF.

## 5.4 Discussion

The purification of functional EcRecN has to date been unsuccessful, despite extensive attempts (Meddows 2002; Neher *et al.* 2006). However, in this chapter the purification of no less than three RecN homologs is described, along with a generalised protocol that could be used to purify others (Figure 5.12). Even though they varied in their physical properties, all three proteins purified in a similar manner. Previously, only the purification of BsRecN has been described, however, the preparations showed chronic GroEL contamination (Sanchez and Alonso 2005). This has also been an issue with the RecN purifications described here, but this protocol does allow exclusion of GroEL and purification of the RecN proteins to homogeneity, as detectable on a polyacrylamide gel.

Due to the structure of SMC-like proteins, with their long coiled-coil arms, the use of gel filtration to identify the size of complexes formed is not ideal, but the suggested sizes are shown. The BfRecN appeared somewhat different, eluting in a smaller volume, suggesting it formed a much larger protein complex (Table 5.3) than the other RecN proteins. It does seem likely that RecN associates with itself to form higher order species, with potential dimers appearing to be stable during SDS-PAGE (Figure 5.11). It is also worth noting that there was a significant difference in the yield of the RecN proteins. Even when expressed in a fermenter, the AqRecN yield was significantly less than the other two, around 1mg of protein per litre of starting culture (Table 5.3). However, good quantities of protein can be obtained by processing larger volumes of induced cell culture. The low level expression of AqRecN in *E. coli* may also explain the inability of this protein to substitute for the native protein *in vivo*. EcRecN is heavily induced during the SOS response (Finch *et al.* 1985b). Therefore, AqRecN may not be expressed to the same level.

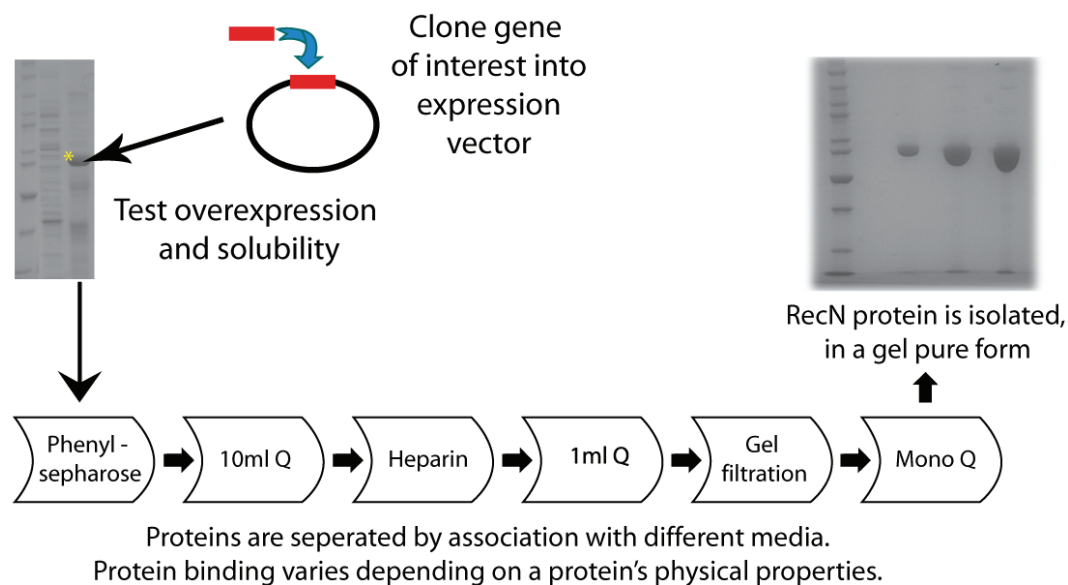


Figure 5.12. Generalised purification scheme for a RecN protein.

The different columns are listed in order in the boxes, RecN containing fractions are eluted, pooled, salt concentration modified and then loaded onto the next column.

**Table 5.3. Differences observed between the RecN homologs during purification.**

Expressed protein	Range protein eluted from S-300 (ml)	Suggested protein size from peak of elution (kDa)	Protein sizes inferred from elution volume (kDa)	Approximate protein yield mg/l of starter culture
<i>H. influenzae</i>	40.0 – 54.1	325	156 – 650	30.00
<i>A. aeolicus</i>	40.2 – 52.3	320	139 – 635	0.75 – 1.00
<i>B. fragilis</i>	37.0 – 44.3	572	401 – 943	27.50

The successful purification of *E. coli* RecN homologs opens up the possibility of studying these proteins *in vitro* and thus to understand how EcRecN, and RecN proteins generally, may function (Chapter 6). It will also allow structural studies, including crystallisation trials, to be commenced. These studies, alongside other *in vitro* assays (Chapter 7), will generate data that will allow a greater understanding of RecN as well as of DSB repair and HR.



# Chapter 6

## The *in vitro* activities of RecN

In the previous chapter I described the purification of three RecN proteins. These include the *H. influenzae* RecN homolog, which crucially functions *in vivo* in place of EcRecN (Chapter 4). Since the insolubility of EcRecN means biochemical characterisation is impossible (Meddows 2002; Nagashima *et al.* 2006), HiRecN provides a viable method to characterise *in vitro* activities of a RecN protein that will complement previous genetic studies. Prior to this, the only reported purification of a RecN protein was that of the *B. subtilis* protein (Kidane *et al.* 2004).

RecN is part of the incredibly diverse SMC family of proteins that have essential roles in many cellular functions, including DNA repair (Sharples and Leach 1995; Hirano 2005b; Hirano 2006; Cortes-Ledesma *et al.* 2007). Despite the array of cellular activities in which SMC proteins are involved, they share many similarities. Principally they form dimers that are integral to larger complexes involved in DNA metabolism. These complexes are generally believed to function by forming stable ring-shaped complexes that can ensnare DNA (Melby *et al.* 1998; Ciosk *et al.* 2000; Haering *et al.* 2002). This entrapment model is based on the activities of cohesin, which is responsible for sister chromatid cohesion. One key question is how does cohesin load onto and entrap DNA? It is believed that cohesin loading is dependent on ATP hydrolysis, which causes conformational changes allowing the complex to temporarily disassociate at the hinge domain. This opens the ring to allow loading onto DNA (Hirano *et al.* 2001; Arumugam *et al.* 2003; Hirano and Hirano 2006; Shintomi and Hirano 2007). This was supported by experiments in which the cohesin complex was expressed as a single, fused Smc1 – Scc1 – Smc3 protein. This fusion complex could still load onto DNA in an ATP dependent manner, as long as the hinge was unperturbed (Gruber *et al.* 2006). However, this is not the only proposed model and more work is needed to confirm SMC functionality (Guacci 2007). Recently work with a related SMC

containing complex, condensin, showed that disruption of the proposed ring complex did not liberate DNA (Hudson *et al.* 2008). It seems plausible, even likely, that mechanisms of SMC-DNA interaction remain undiscovered. RecN is a unique SMC-type protein that can be studied in bacterial systems, with all the inherent benefits of a well-defined model organism. Characterisation of the *in vitro* properties of RecN is therefore important, not only to our understanding of DNA repair, but also to shed further insight into the functions of the SMC protein family.

## **6.1 RecN has a detectable ATPase activity**

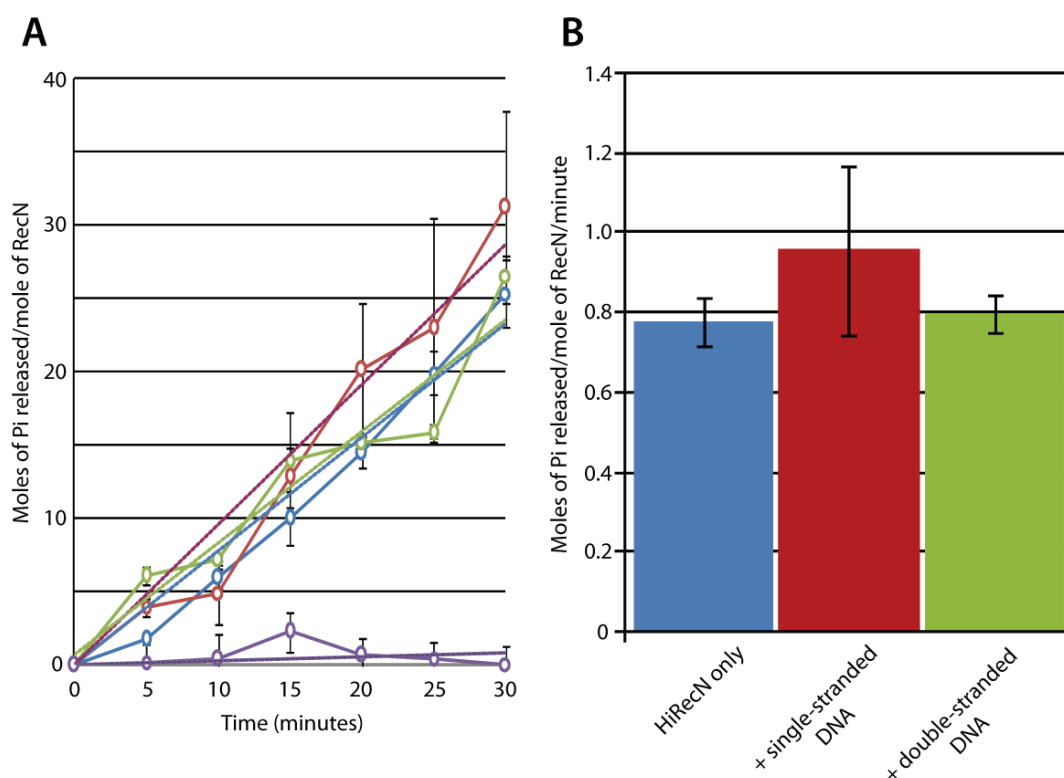
SMC proteins require a functional ATPase activity (Arumugam *et al.*, 2003; Hirano *et al.*, 2001; Hirano & Hirano, 2006; Shintomi & Hirano, 2007). So one of the first biochemical questions asked was whether RecN is an ATPase. When the sequences of several RecN proteins were aligned, there was a high degree of conservation in the head domains, in particular of motifs known to be involved with ATP binding and hydrolysis (Chapter 4). These motifs, known as Walker A and Walker B motifs, are well conserved in a variety of ATPases (Walker *et al.* 1982) and they are both required and intimately involved in the binding and hydrolysis of ATP (Azzaria *et al.* 1989; Loo and Clarke 1995; Szabo *et al.* 1998; Hrycina *et al.* 1999). There is also a third conserved motif unique to the subgroup of ATPases to which SMC proteins belong. This subgroup is referred to as the ABC-type ATPases, as the motifs were initially characterised in ATP binding cassette (ABC)-type membrane transport proteins (Higgins 1992; Holland and Blight 1999; Schmitt and Tampe 2002). However, they have also been characterised in several DNA repair proteins; including UvrA, MutS, and the DSBR protein Rad50 (Aravind *et al.* 1999). The unique third motif of this subgroup is therefore referred to as the ABC signature motif. The signature motif is required along with the Walker A and Walker B motifs to facilitate ATPase activity in this class of protein (Hyde *et al.* 1990). Therefore, the ABC-type ATPases all have a highly conserved set of motifs, which are found in a diverse range of configurations. It is believed that all such ATPases share a common mode of action, whereby ATP binding and

hydrolysis is used to drive conformational changes required for the protein to function (Hopfner and Tainer 2003; Gruber *et al.* 2006).

Studies of the crystal structure of the *Thermatoga* SMC and Rad50 from *Pyrococcus furiosus* showed why ATP binding was essential to allow dimerisation of these SMC proteins (Hopfner *et al.* 2000; Lammens *et al.* 2004). This dimer traps two ATP molecules between the head domains, with the ATP molecules contacting the Walker A and B motifs from one head and the signature motif of the other. However, the gross conformational changes associated with ATP hydrolysis, often termed a “power stroke” (Hopfner and Tainer 2003), have not been observed with SMC proteins (Lammens *et al.* 2004). It has been suggested that the ATP hydrolysis induced conformational changes allows transport of DNA into the ring-shaped cohesin complex, which facilitates sister chromatid cohesion (Weitzer *et al.* 2003). Mutations in the Walker B box that prevent ATP hydrolysis, but not ATP binding, have also been shown to prevent loading of cohesin onto the chromosome (Arumugam *et al.* 2006). BsRecN also has a requirement for ATP; in its absence BsRecN can bind single-stranded DNA, but only in its presence can larger RecN-DNA complexes be formed (Sanchez and Alonso 2005; Sanchez *et al.* 2007a).

Since ATPase activity is vital to SMC proteins, I would expect this to be true for the RecN proteins in this study. To determine if this was the case the ATPase activity of RecN was determined using the malachite green assay (Chapter 2), which quantifies the level of free inorganic phosphate in solution. Detection is achieved by monitoring the colour change of the reagents due to the presence of inorganic phosphate released during hydrolysis of ATP and is quantified by comparison to known standards (Bird *et al.* 1997).

### 6.1.1 RecN is a weak DNA independent ATPase *in vitro*



**Figure 6.1. The *in vitro* ATPase activity of HiRecN.**

All the reactions were performed at 37°C, with HiRecN at 1.5nM reactions were performed in helicase buffer supplemented with 5mM ATP and 5mM MgCl<sub>2</sub>.

- A)** The change in concentration of inorganic phosphate (Pi) present in solution released over time by HiRecN alone (blue), or in the presence of either circular single-stranded  $\phi$ X174 (2 $\mu$ M; red), or circular duplex pBR322, DNA (2 $\mu$ M; green). A no protein control (purple) was also included the error bars represent SEM of four separate repeats for each time point
- B)** The rate of Pi released into solution per molecule of HiRecN alone (blue) and in the presence of either circular single-stranded  $\phi$ X174 DNA (2 $\mu$ M; red), or circular duplex DNA (2 $\mu$ M; green) based on the data gathered for part A, normalised to the no protein control.

Purified HiRecN was shown to turnover ATP, and release inorganic phosphate, at a uniform rate over thirty minutes at 37°C (Figure 6.1 A). The average rate of ATP turnover by HiRecN is approximately one molecule of ATP per minute per RecN molecule ( $0.84 \pm 0.11$  ATP molecules/RecN/minute) (Figure 6.1 B). Although the rate is slightly elevated in the presence of ssDNA,

this increase is not significant and duplex DNA had no apparent affect (Figure 6.1). Similar experiments, using linear duplex and ssDNA, suggested that these DNA species had no affect on the ATPase activity either (data not shown). It appears that HiRecN is a DNA-independent ATPase.

The rate of ATP turnover by HiRecN is consistent with the available data on the ATPase activity of other SMC-type proteins such as cohesin, which has a basal ATP turnover of approximately one ATP molecule/cohesin complex/minute (Arumugam *et al.* 2006) and Rad50, which will turnover approximately 0.2 ATP molecules/Rad50 protein/minute (Hopfner *et al.* 2000). In bacteria the isolated head domain of the *E. coli* SMC homolog, MukB (MukB<sub>342</sub>), was reported to turnover 1.29 molecules of ATP per minute (Lockhart and Kendrick-Jones 1998). However, some other reports conflict with this; the full-length protein was reported to only turnover 0.2 molecules of ATP/MukB dimer/minute (Petrushenko *et al.* 2006a), but a recent study suggested higher rates of ATP turnover of 2-5 molecules of ATP per minute, which increases about two-fold if the complete MukBEF complex is present (Chen *et al.* 2008). Despite the variable rates reported all ATPase activities have comparatively low rates, for instance the ATPase activity of the DNA repair helicase RecG from *E. coli* is in excess of 3000 ATP molecules/RecG molecule/minute (Lloyd and Sharples 1993).

Since SMC proteins associate with DNA, and BsSMC and BsRecN have both been reported to possess DNA-stimulated ATPase activity (Sanchez and Alonso 2005; Hirano and Hirano 2006; Hirano 2006), it would seem contradictory that HiRecN is a DNA-independent ATPase. However, although BsSMC ATPase activity has been shown to be DNA-stimulated, this is not the case with the SMC proteins in the cohesin complex, which are DNA-independent (Arumugam *et al.* 2006). The ATPase activities of SMC proteins appears to be rather complex and although no observable stimulation was seen, it cannot be ruled out that under other conditions, for example in the presence of partner proteins, it may occur.

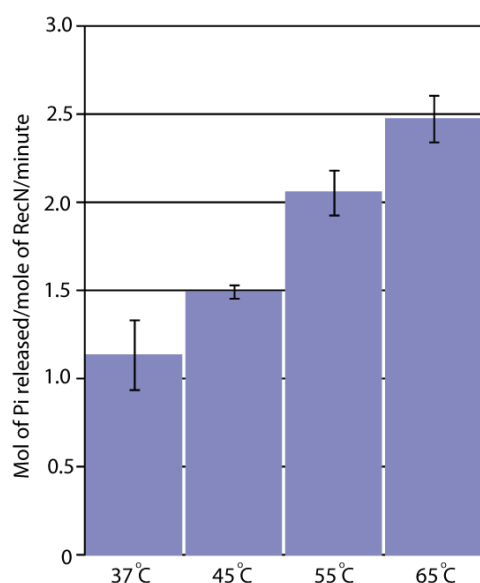
Surprisingly, no ATPase activity could be detected for the *B. fragilis* RecN protein using these assay conditions, the reason for this remains

unknown as RecN from *A. aeolicus* also has ATPase activity (Figure 6.2), which appears to be DNA-independent (data not shown). The reason for my inability to detect any activity for the BfRecN remains unclear although possible explanations are discussed in the discussion of this chapter and could go some way to explaining its inability to function in *E. coli*.

Interestingly, and despite AqRecN being a hyperthermophile, the protein functions at 37°C, with a rate of ATPase activity similar to that of HiRecN at 37°C ( $1.14 \pm 0.20$  ATP molecules/AqRecN molecule/minute). However, it is unclear whether this rate of turnover reflects a true biological activity, as *Aquifex aeolicus* lives at temperatures in excess of 90°C (Deckert *et al.* 1998). Therefore, we investigated the effect of elevated temperatures on the ATPase activity of RecN. According to the  $Q_{10}$  (temperature coefficient) rule, an increase of 10°C should double the rate of a catalytic reaction. In the case of an enzyme this assumes that the increase is within the protein's physiological temperature range and therefore that the enzyme will not be denatured by said increase in temperature, thus HiRecN was not used as trial experiments showed its activity to be diminished at elevated temperatures (>45°C), suggesting the protein is denatured (data not shown).

#### **6.1.2 RecN ATPase activity is not significantly influenced by temperature**

To ascertain if AqRecN ATPase activity was affected by temperature and if its enzyme kinetics obey the  $Q_{10}$  rule, ATPase assays were performed at various temperatures. Since the C-terminal His-tagged AqRecN is easier to purify in larger quantities, it was used instead of the wild-type protein; the purification protocol for AqRecN-His is similar to that described in Chapter 5 but with the addition of a 5ml His-trap HP column (GE Healthcare) as the first purification step. Importantly, at 55°C its ATPase activity was indistinguishable from that of the wild-type protein when assayed in parallel (data not shown) and was approximately two-fold higher than when incubated at 37°C ( $2.06 \pm 0.12$  ATP molecules/RecN/minute, at 55°C) (Figure 6.2).



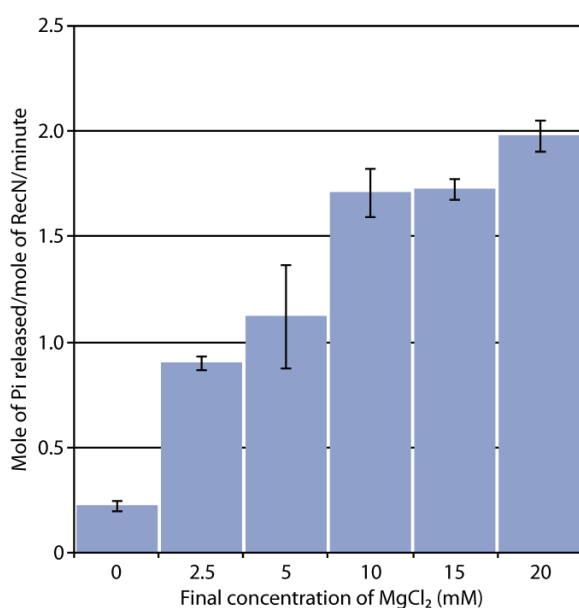
**Figure 6.2. Affect of temperature on the ATPase activity of C-terminal His-tagged AqRecN.**

The protein concentration of AqRecN was kept constant and the reactions incubated for 30min at the temperatures specified in helicase buffer supplemented with 5mM ATP and 5mM MgCl<sub>2</sub>. His-tagged AqRecN was at a final concentration of 2nM. The error bars represent the SEM of four repeats of the experiment at each temperature.

Although it is apparent that increasing the temperature does increase the rate of ATPase activity of His-tagged AqRecN, as expected, it does not obey the  $Q_{10}$  rule. Increases in temperature of 10°C did not double the relative ATPase activity (Figure 6.2 B). However, the increase in rate suggests that the protein is functional and thus stable at these temperatures. Therefore, the rates observed should be comparable. The reason why the ATPase activity does not obey the  $Q_{10}$  rule is unclear, but it is not uncommon in biological systems and may reflect a degree of regulation of the ATPase activity, or the reliance on a set time period for the ATP to be turned over and released, regardless of the temperature. The similar rates of HiRecN and AqRecN at 37°C is surprising. It would be expected that the two proteins would have similar ATPase rates at their physiological temperature if they are to perform the same function. It would seem that the rate of ATPase activity is dependent on temperature and is not regulated *in vitro* so as to be optimal at the organism physiological temperature. Perhaps this represents the basal rate of activity for a RecN at this temperature and *in vivo* other factors can regulate the rate of ATP turnover by RecN.

### 6.1.3 RecN ATPase activity is $\text{Mg}^{2+}$ dependent

A key residue, common to many ABC-type ATPases, is a conserved glutamine residue within the proposed Q-loop motif, which is closely associated with the Walker A motif of RecN. This glutamine residue is involved in the binding of the catalytic  $\text{Mg}^{2+}$  and of the water molecule involved in the nucleophilic attack, making it important for ATPase activity (Rosen *et al.* 1999; Gorbulev *et al.* 2001). The Q-loop forms part of the active site and is a likely candidate for transmitting conformational changes, induced by the binding and hydrolysis of ATP, throughout the molecule (Hopfner and Tainer 2003). However, it has previously been reported that RecN proteins are missing this conserved glutamine residue, it being substituted for a histidine residue in EcRecN (H135) and in HiRecN (H138) (Hopfner *et al.* 2000; Meddows 2002). The substitution of this conserved glutamine led Meddows (2002) to propose that perhaps RecN can bind but not hydrolyse ATP, or alternatively that RecN has an unusual ATPase activity, which does not require  $\text{Mg}^{2+}$  ions (Meddows 2002). The former seems unlikely however, as RecN proteins have detectable ATPase activity *in vitro*. To determine if the ATPase activity of RecN is dependent on  $\text{Mg}^{2+}$  ions, the rate of ATP turnover by HiRecN was measured in the presence of different concentrations of  $\text{MgCl}_2$  over a set time (Figure 6.3).



**Figure 6.3. Effect of  $\text{MgCl}_2$  concentration on the rate of ATP turnover by HiRecN.**

HiRecN at 1.5nM final concentration was incubated at 37°C in helicase buffer, supplemented with 5mM ATP and  $\text{MgCl}_2$  at the final concentrations shown. The quantity of  $\text{P}_i$  released over 30 minutes was measured by the malachite green assay and used to estimate the rate of turnover. The error bars represent the SEM of 4 separate experiments.



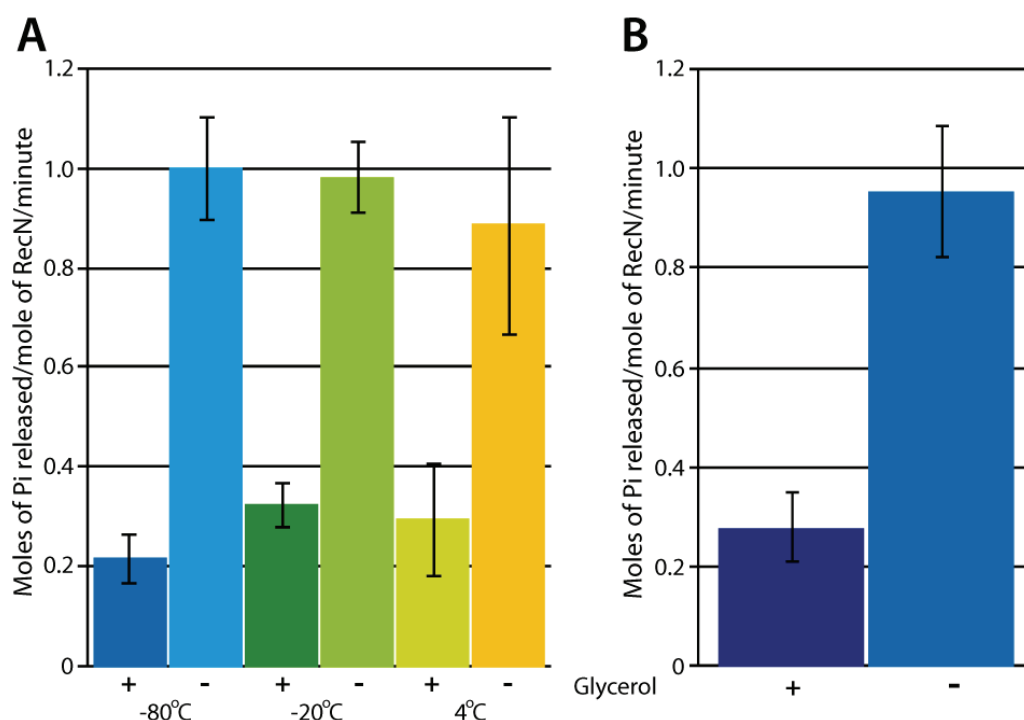
Generally, increasing the concentration of magnesium increased the ATPase activity of the protein, though the effect is reduced as the concentration is increased and the increase appears to reach a plateau. However, the ATPase activity of HiRecN does not appear to be wholly magnesium dependent and some free phosphate was liberated in the absence of added  $\text{MgCl}_2$  (Figure 6.3). This is unlikely to be due to variation in the experiment as no protein and no ATP controls show little variation from zero. However, further analysis is needed to confirm that this is an effect of  $\text{Mg}^{2+}$  ions, such as observing the affect when magnesium acetate is used in place of  $\text{MgCl}_2$ . It has been shown that KCl cannot stimulate the reaction (Dr. Jane Grove, University of Nottingham, personal communication).

These data seem to suggest that HiRecN ATPase activity is largely dependent on magnesium as in its absence the ATPase activity is reduced 5-10 fold. It is possible that the ATPase activity observed when the reaction buffer was not supplemented with  $\text{MgCl}_2$  could be due to magnesium ions present within the protein preparation and carried over from the original cell extract. Unless metal ions are rigorously excluded some carryover is almost unavoidable, especially if they are bound to the target protein.

#### **6.1.4 RecN ATPase activity is affected by glycerol**

Purified RecN proteins stored at 4°C were found to retain their ATPase activity over a period of at least two weeks. However, the activity of samples stored at -80°C was substantially reduced. This could either be as a direct result of the freezing process damaging the protein, or perhaps an affect of the storage buffer. Although the storage buffer was essentially identical to the buffers used during the purification (50mM Tris-HCl pH7.5, 150mM NaCl), it did also include 50% glycerol (v/v), used as a cryoprotectant. I therefore decided to test if the glycerol was affecting the ATPase activity of the protein. Aliquots of HiRecN were diluted with glycerol (to a final concentration of 50% (v/v)) or water to keep buffer concentrations identical. These samples (with and without glycerol) were then either flash frozen in liquid nitrogen and stored at; -80°C or -20°C, or alternatively kept unfrozen at 4°C. After 24 hours

the samples were assayed for ATPase activity, with the frozen samples being rapidly thawed to limit any damage.



**Figure 6.4. Effect of glycerol on ATPase activity of HiRecN**

HiRecN at 15nM was stored overnight in 50mM Tris-HCl pH7.5, with 150mM NaCl and in the presence or absence of 50% glycerol. The protein was diluted 10-fold into the reaction mixture to give a final glycerol concentration of 5%.

- A)** Rate of ATP turnover after storage at the indicated temperature with or without glycerol. The error bars shown SEM of six values for each point.
- B)** The mean of all the reactions with and without glycerol independent of storage temperature. Error bars once again show SEM, but this time of 18 separate experiments for each.

It is readily apparent that, in the presence of glycerol, the ATPase activity of HiRecN is reduced three to four-fold. The reason for this is currently unclear, although it was not an effect of storage temperature (Figure 6.4). However, it does appear to be an inhibitory, rather than a toxic, effect as reduction of the glycerol concentration, by either dilution or dialysis, restores ATPase activity (Dr. Geoff Briggs, University of Nottingham, personal communication). Glycerol is known to affect enzymatic activity (Star activity,

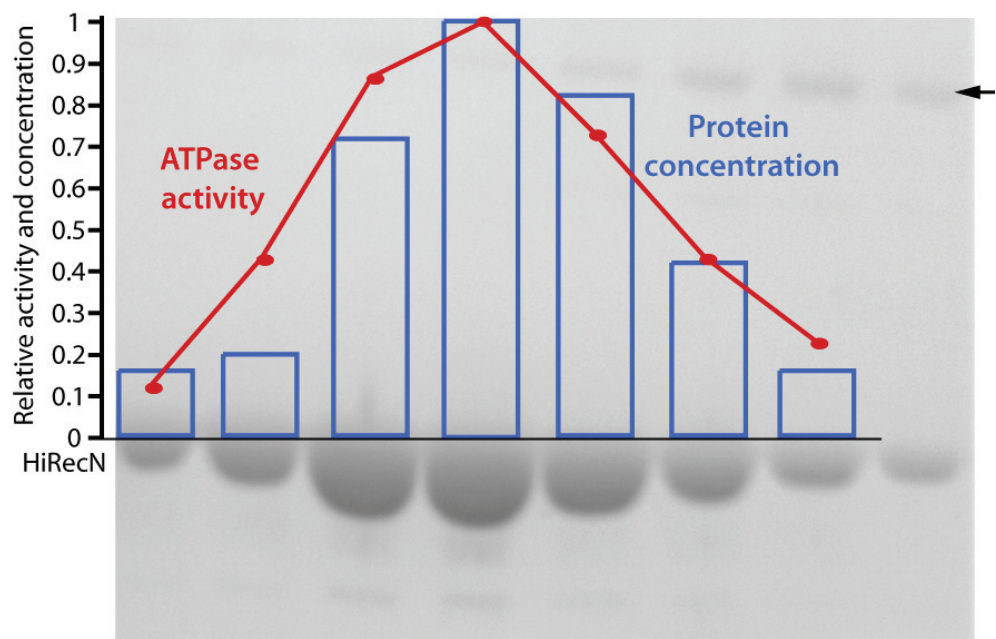
NEB catalog). High concentrations can promote the star activity associated with certain restriction enzymes and there are reports of glycerol affecting the activity of various ATPases. The reasons and effects appear to be varied and complex, but one of the simplest explanations is that the presence of glycerol increases the viscosity of a solution, which has been speculated to inhibit any protein activity that requires conformational changes (Uribe and Sampedro 2003). The greater viscosity of glycerol, compared to water, increases friction between the protein and the solution it is in and thus the conformational changes are limited or slowed, reducing the activity of the protein (Uribe and Sampedro 2003). Glycerol has also been reported to increase the structural rigidity of a protein and this could also restrict conformational changes (Meng *et al.* 2004). Although ATP induced conformational changes have not been demonstrated to occur in RecN, they seem likely, especially when considering that most of the ABC-type ATPases and SMC proteins are believed to undergo gross conformational changes during ATP hydrolysis (Hirano 2002; Hopfner and Tainer 2003; Gruber *et al.* 2006). This provides a plausible explanation for the inhibitory effect of glycerol seen on the proteins' ATPase activity. However, freezing had a minimal affect on the proteins ATPase activity, with rates being comparable between the frozen and unfrozen samples (Figure 6.4), and so it was assumed that the protein could be frozen in small aliquots, in the absence of a cryoprotectant, for later use.

#### **6.1.5 Is the ATPase activity detected truly due to RecN**

The relatively low rate of ATP turnover observed for RecN does raise a problem; it could be that this weak RecN ATPase activity is not due to RecN at all, but is in fact due to the activity of a minor yet highly active ATPase contaminant. The main contaminant that I, and others, have identified when purifying RecN is the chaperone complex GroEL (Kidane *et al.* 2004). The presence of GroEL was not detectable by western blotting, using commercially available antibodies (Sigma), in any of the HiRecN and AqRecN preparations used for this study (data not shown). This suggests that it is not present and it is certainly not a substantial contaminant. However, it is reported to be a major contaminant of the BsRecN preparations used in the studies by Kidane and

Sanchez (2004). The authors of this report argue that as the ATPase activity of GroEL is dependent on  $K^+$  ions (Viitanen *et al.* 1990), GroEL should be inactive under the assay conditions they used. However, analysis of a partially pure preparation of *E. coli* GroEL showed it to have ATPase activity even in the absence of  $K^+$  ions (Dr. Jane Grove, University of Nottingham, personal communication). Also, Viitanen *et al.* (1990) noted that  $K^+$  ions were often present due to carry over from the purification and they had to be 'rigorously excluded'. Kidane *et al.* (2004) assayed BsRecN, purified directly from *B. subtilis* cell, although  $K^+$  ion exclusion limits the ATPase activity of *E. coli* GroEL (Viitanen *et al.* 1990), it may not have the same affect on *B. subtilis* GroEL.

To exclude the possibility that the ATPase activity I detected is due to a contaminant, two approaches were taken. Firstly, individual fractions containing HiRecN, eluted from the gel filtration column during the purification (Section 5.4), were assayed for ATPase activity and total protein concentration was measured. These samples were not yet completely clean and minor contaminants are clearly visible upon SDS-PAGE analysis (Figure 6.5). What is interesting is that the ATPase activity corresponds to the total protein concentration and its peak coincides with the highest concentrations of HiRecN. There is a minor contaminant visible upon SDS-PAGE analysis and peak elution of this protein occurs after that of the HiRecN, but it does not influence the ATPase activity. This would strongly suggest that the ATPase activity detected is not due to the presence of a contaminant but to the HiRecN. However, this does not exclude the possible presence of a co-eluting contaminant that may not be visible on SDS-PAGE analysis, causing the ATPase activity observed. So a second approach was taken, creating ATPase deficient RecN mutants and assaying their ATPase activity.



**Figure 6.5 Relative ATPase and protein concentration of fractions off a gel filtration column of HiRecN.**

The relative, ATPase activity is indicated by the red line and total protein concentration by the blue bars. The graph is overlaid onto an SDS-PAGE of the fractions obtained from the gel-filtration column to show the protein concentrations. The arrow indicates a contaminant present in some of the fractions.

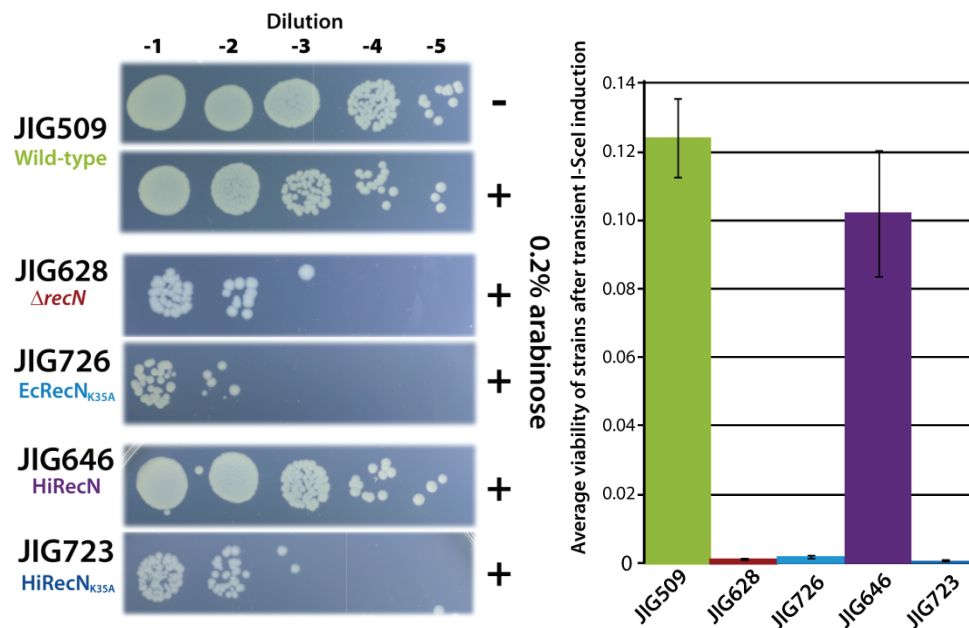
## 6.2 RecN requires ATPase activity to function

The definitive method to confirm that the ATPase activity detected is due to RecN, and not a contaminant, is to generate and purify a RecN mutant that is deficient for ATP hydrolysis. To do this a mutation in the Walker A motif was made, substituting the highly conserved lysine residue (at position 35) for alanine (Walker A motif; GXXXXK/T/S). Numerous studies have shown that mutation of this residue inhibits the ATPase activity of ABC-type ATPases. This inhibition is believed to be as a result of mutants being unable to bind ATP (Rozen *et al.* 1989; Schneider and Hunke 1998; Lapinski *et al.* 2001; Frelet and Klein 2006). Therefore, the RecN<sub>K35A</sub> mutant proteins should be ATPase deficient. Furthermore, it will be informative to determine whether the ATPase mutation results in a detectable phenotype when introduced into the

chromosome of *E. coli*, and if so, how similar this phenotype is to that of a  $\Delta recN$  strain.

The *A. aeolicus* *H. influenzae* and *E. coli* *recN* gene sequences were mutated and cloned into the same plasmids used for overexpression (Chapter 2 and Table 5.2) and the mutant RecN proteins were overexpressed and purified in the same manner as the wild-type proteins. Like EcRecN, EcRecN<sub>K35A</sub> was present in the insoluble fraction and so could not be studied. The HiRecN<sub>K35A</sub> and the His-tagged AqRecN<sub>K35A</sub> proteins were soluble and purified in the same manner as the wild-type proteins. The malachite green ATPase assay revealed free Pi at only background levels, with no detectable increase over time. This strongly suggests that the K35A mutation eliminates the ATPase activities of these two proteins and that the previously detected ATPase activities were due to the RecN proteins and not due to a contaminant. RecN is therefore an ATPase. I proceeded to investigate if the ATPase activity is essential to the function of RecN *in vivo*.

The genes encoding EcRecN<sub>K35A</sub> and HiRecN<sub>K35A</sub> were introduced into the *E. coli* chromosome, replacing the EcRecN coding sequence and leaving the promoter intact, as was described previously for the creation of the heterologous *E. coli* strains carrying foreign *recN* coding sequences (Chapter 4). The presence of the mutations was confirmed by sequencing and the mutations were then moved into the I-SceI expression and cleavage strain and the viability of *E. coli* carrying these mutations assayed by exposure to I-SceI. It is clear that the K35A mutation in both proteins makes the cells as sensitive to DSB formation as a  $\Delta recN$  (Figure 6.6). It would appear that loss of the ATPase activity results in a complete loss of function of the RecN proteins *in vivo*.



**Figure 6.6. The viability of *E. coli* strains expressing ATPase deficient RecN proteins, after I-SceI induced formation of DSBs.**

- A) Serial dilution and plating of heterologous RecN strains carrying the I-SceI expression cassette and cleavage site after being induced to express I-SceI by exposure to arabinose.
- B) Graph showing the average viability of such strains after transient I-SceI induction compared to the same cultures grown in glucose. Error bars represent the SEM of four separate experiments, except for JIG509 which was repeated seven times

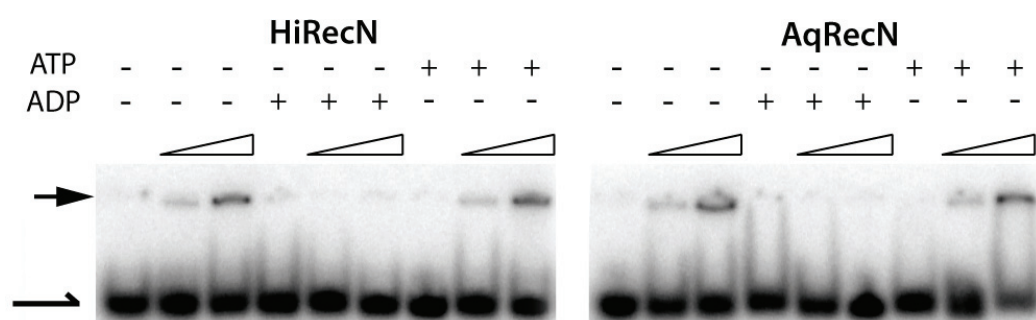
### 6.3 DNA Binding Activities of RecN

SMC proteins are almost universally associated with DNA and this is also the case with several other ABC-type ATPases (Hopfner and Tainer 2003; Hirano 2006). RecN is an SOS induced protein, known to be involved in recombination and DNA repair (Picksley *et al.* 1984a; Finch *et al.* 1985b) and in *B. subtilis* the RecN protein has been shown to bind and aggregate ssDNA (Sanchez and Alonso 2005; Sanchez *et al.* 2007a). Therefore, an interaction of the RecN proteins purified in this study with DNA would seem to be likely. Although BsRecN has been reported to only interact with ssDNA, SMC

proteins have been reported to interact with a variety of DNA substrates; including single-stranded and duplex DNA (Hirano and Hirano 2004; Hirano 2005b). The effect of nucleotides was also investigated, since both the BsSMC and BsRecN proteins can interact with DNA in the absence of ATP, but in its presence they cause aggregation of the ssDNA (Hirano and Hirano 1998a; Hirano *et al.* 2001; Sanchez and Alonso 2005; Sanchez *et al.* 2007a). Therefore, gel retardation assays were used to detect any possible interactions with a variety of DNA substrates and a variety of nucleotides were added to detect any effect on the gel retardation assays.

Since glycerol has been shown to have inhibitory affect on RecN's ATPase activity, and could inhibit other activities (Figure 6.4), glycerol-free helicase buffer (Chapter 2.1.4) was used in the gel retardation assays. All the retardation experiments were performed as described in Chapter 2, with the reactions being incubated for 30 minutes at either 37°C for HiRecN, or 55°C for AqRecN, before loading. No difference was seen between using a glycerol or ficoll based loading buffer, which was added immediately prior to loading to ensure the samples were dense enough to enter the wells. This is despite the suspected inhibitory effect of glycerol on the ATPase activity of RecN protein.

### 6.3.1 RecN binding to linear ssDNA is ATP independent



**Figure 6.7. Gel retardation of linear ssDNA by RecN proteins.**

The effect of HiRecN and AqRecN, at final concentrations of 0nM, 250nM and 500nM, on ssDNA (50-mer, RGL16, 2nM final concentration the arrow head shows the 3' end) is shown. The protein was pre-incubated for 30min with the labelled DNA in helicase buffer supplemented with MgCl<sub>2</sub> (5mM final concentration) and nucleotides as indicated (1mM final concentration). The arrow indicates the wells where the retarded material can be seen.



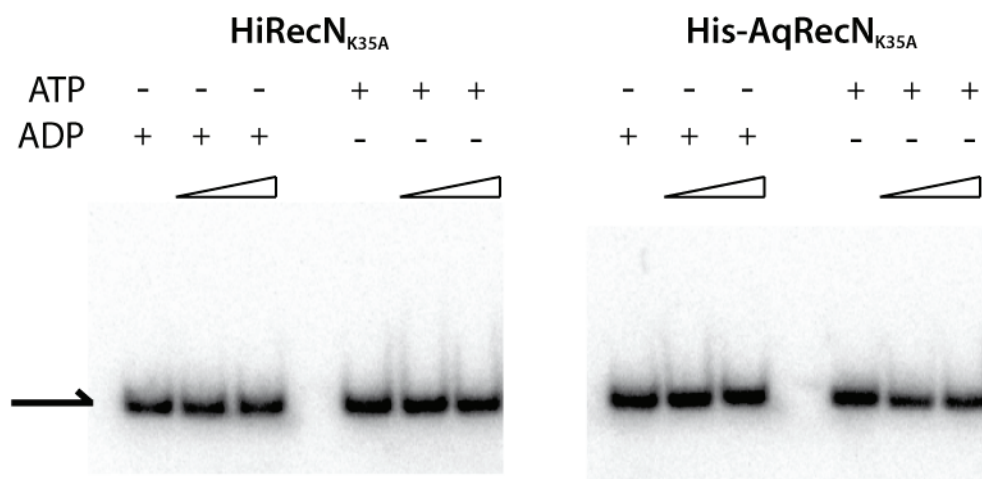
Despite extensive attempts, DNA binding was not seen with either of the RecN proteins. However, a small quantity of labelled material was consistently observed in the wells, when the protein concentration is at over a 100-fold excess compared to the labelled oligonucleotide (Figure 6.7). The retention of labelled material within the well would suggest that it is part of a very large protein-DNA complex, which cannot migrate into the gel. Interestingly, this interaction is inhibited in the presence of ADP, with no labelled material being retained in the well. The same effect was observed with both HiRecN and AqRecN. The retention of material in the well seems to represent some form of RecN-DNA interaction, although it is not a stochastic binding event. Therefore, it appears that under the conditions tested HiRecN and AqRecN are not DNA binding proteins.

Perhaps RecN works like the SMC complexes to entrap DNA. Therefore, it could potentially trap multiple DNA molecules resulting in large DNA-protein aggregates that cannot migrate into the gel. Rad50 and ABC-type transporter proteins make use of ATP binding and hydrolysis to drive the conformational changes that facilitate their activity (Hopfner *et al.* 2000; Smith *et al.* 2002). Therefore, it is possible that ADP induces conformational changes in the RecN protein that prevent it interacting with DNA. The turnover of ATP also drives the conformational changes of the SMC proteins in the cohesin complex. These changes allow DNA to be transported into the interior of the ring-complex, causing entrapment of the DNA (Gruber *et al.* 2006). However, BsSMC can interact with DNA in a nucleotide independent manner, although ATP hydrolysis can drive other conformational changes that allow the assembly of large BsSMC-DNA complexes. Therefore, it would appear the affect of ATP hydrolysis driven conformation changes varies between SMC proteins, but it is generally required for their ability to interact with DNA.

However, it seems likely that ATP hydrolysis is driving conformational changes in RecN. Perhaps in the presence of ATP, or no nucleotide, the protein can assume a conformation that allows it to interact with DNA. However, in the presence of ADP the protein is fixed in a release conformation preventing

the DNA interaction. The hydrolysis of ATP could therefore allow RecN to release DNA in a regulated manner.

Next, I investigated whether the RecN K35A mutant proteins could still interact with DNA. The inability of these mutations to function *in vivo*, or as ATPases *in vitro*, has already been discussed. When the K35A mutant proteins of HiRecN and AqRecN were incubated with ssDNA, radiolabeled material was not retained in the wells (Figure 6.9). Surprisingly, proteins with the K35A mutation do not appear to interact with DNA, even in the limited manner seen with the wild-type proteins. Since the RecN proteins did not need to bind ATP to allow the RecN-DNA interaction it could be assumed that preventing ATP binding would have a minimal effect on this activity. However, this is not the case. It is possible that the mutation has induced more dramatic affect on the structure of the protein than expected preventing its activity. This does seem unlikely since the mutant proteins purified in a manner almost identical to that of the wild-type proteins. Alternatively the residue may have other roles in the function of the protein and the loss of DNA binding is a secondary effect, not directly connected to the loss of ATP binding.

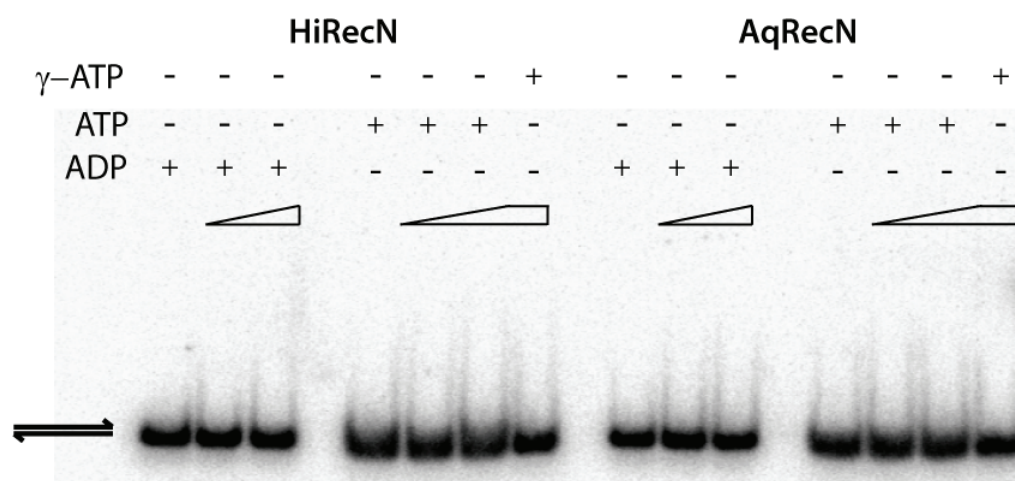


**Figure 6.8. Gel retardation of single stranded DNA by RecN proteins carrying the K35A ATPase mutation.**

The effect of the mutant HiRecN and his-tagged AqRecN are shown. Both were pre-incubated with the ssDNA (50mer, RGL16 arrow head denotes the 3' end, 2nM) in Helicase buffer supplemented with MgCl<sub>2</sub> (5mM, final concentration) and nucleotides as shown (1mM final concentration), protein concentrations of 0nM, 250nM and 500nM were used. DNA was at a final concentration of 2nM.

### 6.3.2 The DNA binding activity of RecN is specific to ssDNA

The SMC proteins are reported to interact with a variety of DNA substrates, including duplex DNA (Hirano and Hirano 2004). Therefore, further experiments tested whether the RecN proteins were able to interact with double-stranded DNA substrates. Linear duplex DNA was made by annealing a complementary sequence to the single-stranded oligonucleotide used previously. No DNA binding or well retention was observed when this oligonucleotide was incubated with the purified RecN proteins (Figure 6.9). Attempts were also made using the non-hydrolysable Gamma – S – ATP to see if this could induce RecN to remain in a binding receptive form, all of these assays failed to show any interaction between RecN and the duplex oligo. Therefore, it would appear that RecN is unable to interact with this duplex DNA.

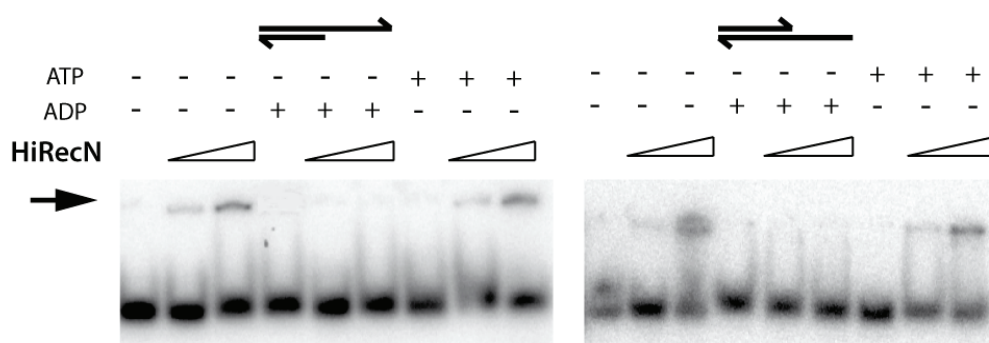


**Figure 6.9. Gel retardation of duplex DNA by RecN proteins.**

The affect of HiRecN and AqRecN at final concentrations of 0nM, 250nM and 500nM, are shown. Both proteins were pre-incubated with the linear-duplex 50mer DNA (RGL16 annealed with ELB37, 2nM final concentration, wrrow heads show the 3' ends) in Helicase buffer supplemented with  $MgCl_2$  (5mM final concentration) and nucleotides as indicated (1mM final concentration).

Since BsRecN has been reported to bind ssDNA preferentially that with an exposed 3'-OH group (Sanchez and Alonso 2005; Sanchez *et al.* 2007a), the interaction of HiRecN with DNA structures containing 3' or 5' overhangs was

investigated. It is possible that RecN will show a preference for structures, which contain a duplex/single-stranded DNA interface. These could potentially mimic the DNA structures found at a break site after end processing and could indicate if RecN shows a preference for 5' or 3' ssDNA ends. Overhangs were created by annealing complementary 25 and 50 oligonucleotides to give a 3' 25 oligonucleotide or 5' 25 oligonucleotide overhang. HiRecN interacts with both types of DNA overhang in exactly the same manner, resulting in the retention of a small quantity of labelled DNA within the well. This retention of material in the well is inhibited in the presence of ADP (Figure 6.10). These results are almost identical to those obtained for HiRecN incubated with the single-stranded 60 oligonucleotide, suggesting that HiRecN does not bind to a duplex/single-stranded DNA interface, but is only binding to the single-stranded region. The results also show that the binding of ssDNA is not affected by the presence of either 3' or 5' ends.



**Figure 6.10. Gel retardation of DNA fragments with single stranded overhangs, by HiRecN proteins.**

The effect of HiRecN, at a final concentration of 0nM, 250nM and 500nM on two types of DNA overhang is shown (both at a final concentration of 2nM arrow heads show the 3' ends). The HiRecN was pre-incubated for 30min with labelled DNA in helicase buffer supplemented in with  $MgCl_2$  (5mM final concentration) and nucleotides as indicated (1mM final concentration). The arrow indicates the wells where the retarded material can be seen.

As both RecN proteins show a preference for ssDNA the effect of HiRecN on circular, rather than linear ssDNA was also tested. If the RecN protein functions like a cohesin, trapping DNA within its structure, then its

affect on circular DNA should be more pronounced than that on linear, simply because the DNA cannot ‘fall out’ of the ring complex. Therefore, gel retardation assays using HiRecN and circular single-stranded  $\Phi$ X174 DNA were performed. The results were inconclusive, although in the presence of the wild-type protein a small quantity of material appeared to be retarded (data not shown). It could be that HiRecN is unable to interact productively with circular ssDNA in the same way as it does the linear substrate. However, further experiments will be required to establish whether this is indeed the case.

## **6.4 Structural properties of RecN**

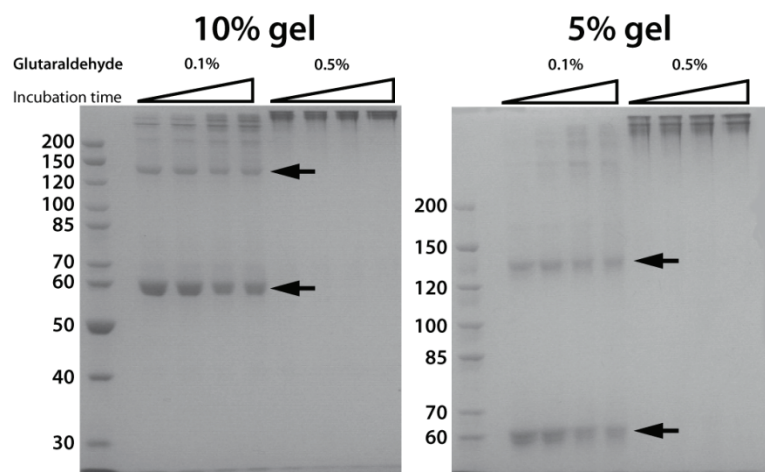
The sequence and predicted structural homology of RecN to SMC proteins led to the speculation that RecN was a member of this family of proteins. These studies have however concentrated on sequence analysis, in particular the conserved Walker A and B boxes and the Q-loop (Sharples and Leach 1995; Meddows 2002). I have already presented some data on the sequence conservation between RecNs from different species (Chapter 4). Here the tertiary and quaternary structure of RecN is investigated, in particular the possibility that RecN proteins may form functional complexes.

### **6.4.1 RecN can form multimer complexes *in vitro***

It is almost universal that SMC proteins form dimers and complex with other proteins. Even the BsSMC, which can function as a simple homodimer, has partner proteins that modulate its activity (Hirano 2002; Hirano 2006). The available data on RecN from *B. subtilis* suggests that while it associates with DNA on its own, it does form large protein-DNA complexes (Sanchez and Alonso 2005; Sanchez *et al.* 2007a). It would therefore seem likely that RecN would also dimerise or multimerise and perhaps interact with other proteins as well. During purification it was noted that that HiRecN, AqRecN and BfRecN proteins eluted from a gel filtration column in a manner that suggested their apparent size was much greater than 60kDa (Chapter 5). A similar observation of BsRecN suggested it eluted as an octameric complex (Kidane *et al.* 2004).

However, size estimates by gel filtration assume the protein is spherical. As stated before, being an SMC-like protein, RecN is unlikely to be a spherical molecule and due to this will have a larger Stokes radius than its actual size suggests. Although the Stokes radius does have an effect, it is unlikely to lead to the increases in apparent size observed alone, as some estimates suggested that RecN was running as a complex in excess of 600kDa in size. It was also noted during the purification that if excessive quantities of pure protein were loaded onto a gel a faint band, approximately twice the size of the three RecN proteins purified in this study, became visible. Sequencing suggested that this was RecN and so it was proposed that rather than this band being a contaminant, it actually represented a dimeric species of RecN that could form even under the denaturing conditions used (Chapter 5). The RecN proteins also have a propensity to aggregate both *in vitro* and *in vivo* and this could lead to the large protein complexes observed during purification of EcRecN (Meddows 2002; Nagashima *et al.* 2006). It would seem likely therefore that the RecN proteins can interact with themselves forming at least dimeric and possibly even larger complexes, although, whether this represents a coordinated complex formation, or random aggregation, is unclear.

To investigate these possibilities, glutaraldehyde cross-linking was used to detect if any higher order RecN species existed, even transiently, in solution. Exposure to glutaraldehyde will covalently link molecules that are in close proximity. As a result molecules become permanently linked and the complex formed can then be identified by SDS-PAGE under denaturing conditions. The number of bonds formed depends upon the concentration of both the protein and glutaraldehyde used. Purified AqRecN was incubated for 30 minutes with glutaraldehyde and then analysed by SDS-PAGE (Figure 6.13).



**Figure 6.11. Affect of glutaraldehyde cross-linking on the migration of AqRecN.**

AqRecN at 0.25mg/ml was cross-linked by exposure to glutaraldehyde at 0.1% or 0.5% final concentration, for 5, 10, 20 or 30 minutes at 4°C and then analysed by SDS-PAGE. Gels of different concentrations were used to allow size determination of the species present. The arrow indicates the position of the monomeric and dimeric RecN species.

In the absence of glutaraldehyde all the AqRecN protein is visible as a monomer, running as a single band of protein a little under 60kDa in size. Similarly, at low glutaraldehyde concentrations the majority on the AqRecN remained as a monomer (Figure 6.11). However, a larger band of approximately a 120kDa is also clearly visible, especially on the 5% gel. This makes it tempting to speculate that a dimer of RecN exists in solution. There are also much larger RecN complexes visible on the gel and these may reflect random cross-linking events or perhaps the functional form of the protein. Similar results were obtained when HiRecN was cross-linked with glutaraldehyde, with both dimeric and larger RecN species detectable (Dr. Jane Grove, University of Nottingham, personal communication).

The larger RecN complexes became prevalent at increased (0.5%) glutaraldehyde concentration. As cross-linking is random and will link any protein in close proximity there will always be a certain percentage of random cross-linking of several RecN monomers which would give the larger complexes seen here. Interestingly on the 5% gel they can be seen to form several discrete bands in excess of 200kDa. These complexes are largely absent at the 0.1% glutaraldehyde concentration and those that are present

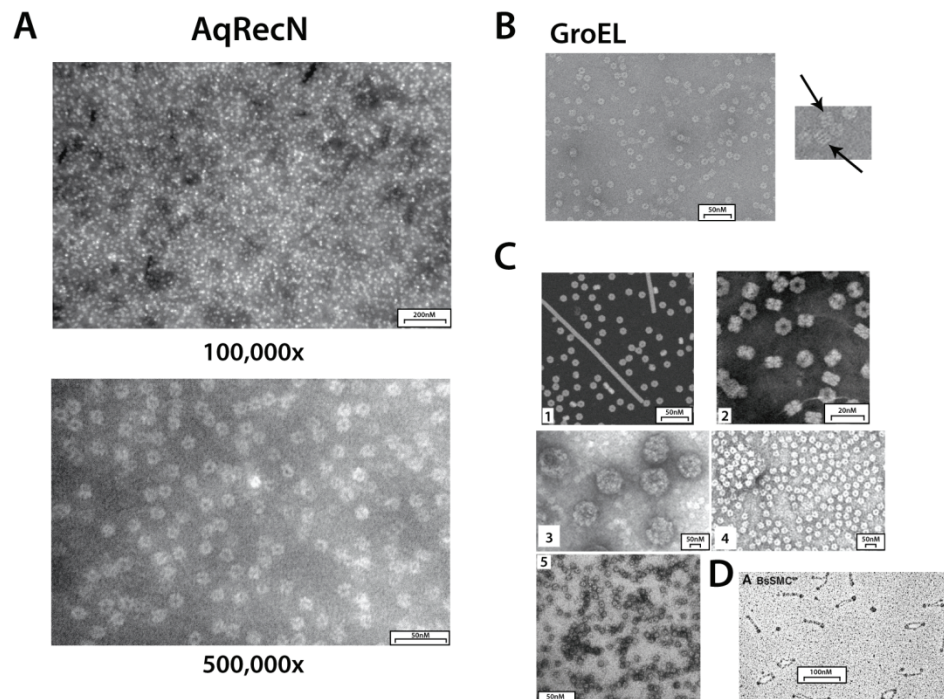
appear to be smaller (Figure 6.13). Whether these bands represent distinct species of RecN formed into large protein networks, as suggested by Sanchez and Alonso (2005), or random cross-linking events of the proteins remains ambiguous at this stage. The previously mentioned tendency of RecN protein to aggregate could also affect these results, increasing the chance of cross-linking numerous RecN molecules.

#### **6.4.3 AqRecN forms discrete complexes as visualised by EM**

Electron microscopy of SMC proteins revealed that the dimers appear as distinctive V-shaped structures, with the two SMC proteins joined at their hinge region (Melby *et al.* 1998; Anderson *et al.* 2002; Haering *et al.* 2002). While RecN is much smaller, the BsRecN protein has also been visualised by AFM, being identified as ‘discrete aggregates’ on DNA (Sanchez *et al.* 2007a). It should therefore be possible to visualise HiRecN and AqRecN, and possibly identify any structures formed in solution.

Purified AqRecN and HiRecN were therefore analysed using transmission electron microscope (TEM), specifically a JEOL JEM 1010 TEM at 100kV at magnifications ranging from 100 – 500k. Images captured of AqRecN reveal discrete structures that are present throughout the visual field (Figure 6.14 A). These structures appear to be ring-shaped with a central cavity, whether this cavity is a hole, or merely an indentation, is unclear. Interestingly, with an approximate diameter of 5-10nm the rings are roughly the correct size to be the heads of a RecN dimer assuming they are a similar size to those of SMC proteins (Melby *et al.* 1998). Closer examination shows a non-uniform, almost bumpy or spiky, edge to the protein complexes, perhaps due to protrusions from the main molecule. One possible explanation for this structure is that the globular heads of two or more RecN proteins form the ring with the long coiled-coil region jutting away from the molecule giving rise to the protrusions.





**Figure 6.12. Transmission electron microscopy images taken of AqRecN and comparison to other proteins including SMC's.**

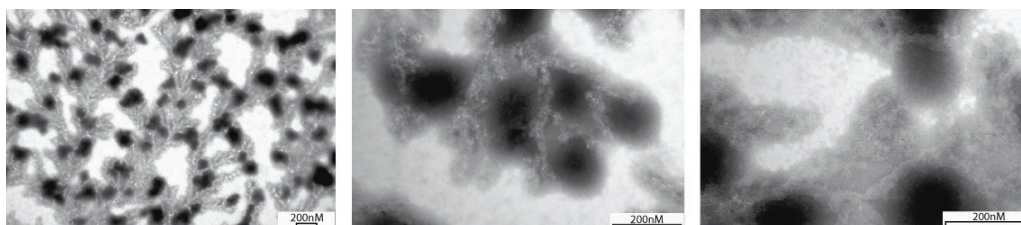
The scale is shown on each picture; all the proteins are negatively stained although various stains were used.

- A)** AqRecN protein shown at two magnifications, negatively stained with 1% PTA. These images were taken with the kind assistance of Dr. Katy Evans (University of Nottingham).
- B)** GroEL image obtained from [www.planetesacha.com](http://www.planetesacha.com), the side-on images of GroEL are highlighted by arrows.
- C)** Shows EM images of three other proteins, C<sub>1</sub> and C<sub>2</sub> are Earth worm haemoglobin (images obtained from [www.biology.bnl.gov](http://www.biology.bnl.gov) shown at two magnifications. The long rod shaped objects are of the tobacco mosaic virus particles, used as internal controls. C<sub>3</sub> and C<sub>4</sub> show EM images of virus capsid protein, C<sub>3</sub> is largely comprised of assembled protein in large spherical objects but smaller particles are visible in the background, C<sub>4</sub> is a mutant form of the protein that cannot assemble and remains monomeric, both were obtained from Bishop *et al.*, (2007). C<sub>5</sub> shows Haemoporphin, one of the major components' of the haemolymph of *Aplysia californica*, obtained from Jaenicke *et al.*, (2003). The protein is believed to allow the formation of membrane channels (Jaenicke *et al.* 2003).
- D)** Image of BsSMC showing the characteristic structure of known SMC proteins under EM (Melby *et al.* 1998).

The structure observed for AqRecN is however very similar to that of GroEL, the oft-mentioned contaminant common in RecN preparations (Sanchez and Alonso 2005) (Figure 6.14 B). However, the protein was confirmed by MALDI-PMF sequencing. Furthermore, when viewed from the side the GroEL molecule appears to be a much longer barrel-shaped structure that has a distinct banded patterning and only end-on does it appear as a ring (Figure 6.14 B). No barrel-like structures are visible in the images of AqRecN (Figure 6.14 A).

Ring-shaped structures are relatively common to several known protein complexes (Figure 6.14 C). In Figure 6.14 C, there are three sets of electron micrographs of different, and unrelated proteins. Frames C<sub>1</sub> and C<sub>2</sub> show earthworm haemoglobin, a globular molecule that resembles both GroEL and AqRecN (Figure 6.14 C<sub>1</sub>), but on closer examination the elongated barrels are not visible. Instead it forms a more globular molecule that lacks the distinguishing banding pattern of GroEL (Figure 6.14 C<sub>2</sub>). The next two frames show images of the L1 capsid protein of human papilloma virus (Figure 6.14 C<sub>3</sub> and C<sub>4</sub>). This protein assembles into a large spherical structure, the virus capsid, while in the background smaller doughnut shaped molecules can be seen. These doughnut shaped molecules are unincorporated monomers of the capsid proteins (Figure 6.14 C<sub>3</sub>). When a mutant of this protein, which cannot assemble into the capsid, was visualised the individual capsid proteins are more readily visible (Bishop *et al.* 2007). These capsid monomers also formed discrete doughnut-shaped structures with spiky edges, much like AqRecN (Figure 6.14 C<sub>4</sub>). Finally, the haemoporphyrin protein shown in frame C<sub>5</sub>, also forms a doughnut-shaped molecule, again being somewhat spiky in appearance and of a similar size to the structures seen for AqRecN. Interestingly, these molecules are all unrelated and vary in size, yet they form remarkably similar shapes, as revealed by electron microscopy (Figure 6.14) (Jaenicke *et al.* 2003; Bishop *et al.* 2007) and this may reflect a tendency for certain gross structures to look similar. These observations support the suggestion that AqRecN forms a complex, although its appearance is quite distinct from that of SMCs.

Surprisingly, such structures were not visible when the HiRecN was examined. Instead the protein appears to be amorphous in nature and no distinct structures could be observed (Figure 6.15)

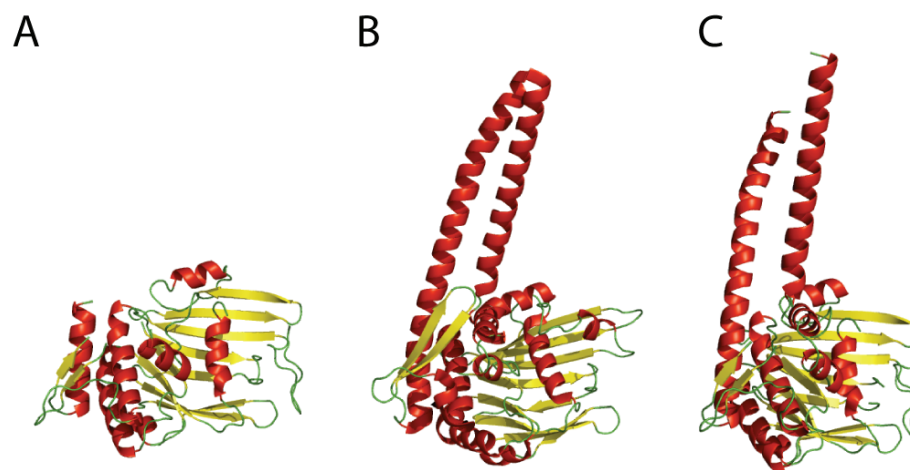


**Figure 6.13. Transmission electron microscopy images of HiRecN.**

The protein is at a concentration of approximately 1mg/ml and is negatively stained with 1% PTA. Magnification increases as you go from left to right and scale is shown by the bars. Images produced with the kind assistance of Dr. Katy Evans (University of Nottingham).

#### **6.4.4 Model of RecN structure provides evidence for dimerisation**

The resolution of crystal structures has proven to be extremely informative, providing information on both the functions and interactions of a protein. However, there is no crystal structure for a RecN protein, or even of a full-length SMC family member (Hopfner and Tainer 2003; Lammens *et al.* 2004). There have, however, been successful attempts to crystallise parts of SMC proteins, particularly the head domains, and the crystal structure of several SMC head domains are available. This includes the head domains of *P. furiosus* Rad50, the *T. maritima* SMC and the *E. coli* SMC homolog MukB (van den Ent *et al.* 1999; Hopfner *et al.* 2000; Lowe *et al.* 2001). Using the data from the first two structures, Dr. Geoff Briggs (University of Nottingham) created a model of RecN (Figure 6.16). The probability Z-score of the model protein suggested a high likelihood of the model being accurate. Despite the structure being limited to the head domain, the model does provide a way to understand the activity of RecN, as well as providing insight into how site-directed mutagenesis might prove informative.

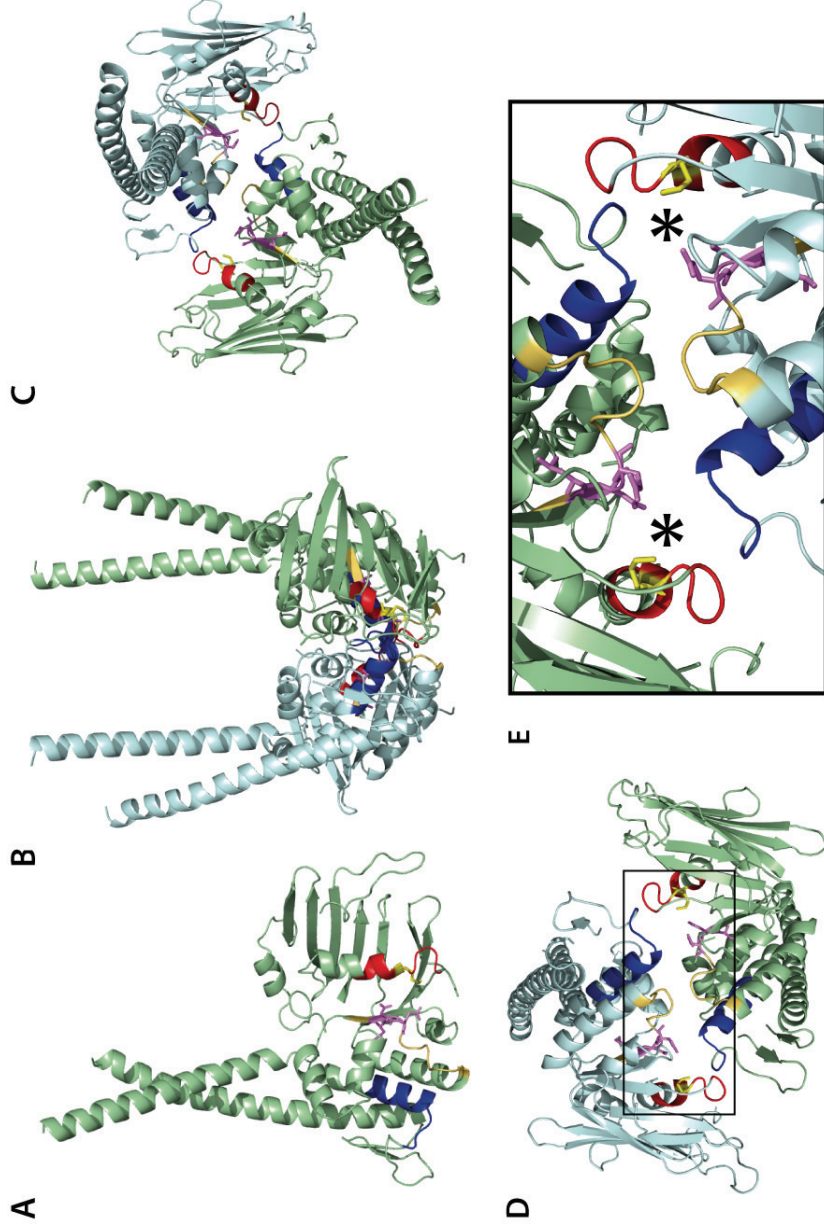


**Figure 6.14. Construction of a model of *E. coli* RecN.**

The secondary structure including  $\beta$ -sheets (yellow), flexible loops (green) and  $\alpha$ -helices (red).

- A)** Structure of *T. maritima* SMC head domain (Lowe *et al.* 2001).
- B)** Structure of *P. furiosus* Rad50 head domain (Hopfner *et al.* 2000).
- C)** Modelled structure of *E. coli* RecN. This was created via sequence alignment of EcRecN using ClustalW v1.83 and applied to the other structures using Chainsaw, a component of the CCP4 suite. Images were taken from PyMOL for Windows (www.pymol.org).

The model was used to help understand how RecN may function. AqRecN and HiRecN have been shown to be ATPases. When the model is examined it can be seen that although the Walker A and B motifs are in close proximity and could form a binding pocket, however, the signature motif can be seen to be spatially distinct (Figure 6.17 A). It has been suggested that SMC proteins dimerise to allow the signature motif from one head to contact and interact with the Walker motifs of the other (Lammens *et al.* 2004). This occurs in a reciprocal manner, thus forming two ATP binding pockets and allows the heads come together to create a nucleotide-sandwiched dimer (Figure 6.17 E) (Haering *et al.* 2004; Lammens *et al.* 2004). A similar arrangement has been postulated in other ABC-type ATPases (Hopfner *et al.* 2000; Lammens *et al.* 2004). This model would also be consistent with the dimeric complexes detected by cross-linking, with the heads providing at least one potential dimer interface (Figure 6.13).



**Figure 6.15. Further analysis of the model structure of *E. coli* RecN.**

The model was produced as described in Figure 6.16 and the dimer made by manually docking the model structures. Key conserved domains associated with ATPase activity are highlighted, the signature motif (dark blue), the Walker A box (red) and the Walker B box (gold). Key residues involved in ATP hydrolysis are highlighted separately, and shown as stick diagrams. This includes the K35 residue (yellow) within the Walker A motif, which was mutated as part of this study and the FDE residues of the Walker B box (purple)

- A) Side on view of the RecN model, showing the three conserved domains of interest.
- B) Potential RecN dimer showing association of the head domains and ATPase motifs.
- C) Top view of the potential dimer.
- D) View looking up into the dimer.
- E) Enlarged view showing the potential interaction of the domains on the two RecN molecules allowing the creation of two active ATP binding and hydrolysis sites (indicated by asterixes).

## 6.5 Discussion

Since the discovery 25 years ago of RecN as a component of the SOS response associated with the repair of DSBs, a wealth of genetic evidence about the function of the protein has been described (Lloyd *et al.* 1983; Sargentini and Smith 1983; Finch *et al.* 1985b; Sharples and Leach 1995). However, even with the recent reports on the biochemical activities of BsRecN, there is still much we do not understand about its enzymology and we have few clues as to how RecN is involved in the repair of a DSB. In this chapter, biochemical data is presented for two RecN proteins, along with mutational analyses of both.

HiRecN and AqRecN have a weak ATPase activity, turning over roughly one molecule of ATP per minute per molecule of RecN. I have shown this ATPase activity to be essential for the repair of DSBs, as the K35A mutation of EcRecN and HiRecN, eliminates its ATPase activity *in vitro* and results in sensitivity to DSBs *in vivo*. The reliance on active ATPase activity for the repair of DSBs could also explain the inability of BfRecN to function in *E. coli*. This RecN has no detectable ATPase activity under the conditions used and one could infer that perhaps it cannot function as an ATPase within *E. coli*. Whether BfRecN can act as an ATPase under the correct experimental conditions is unclear, as they were not found in this study. I would favour the idea that BfRecN is an ATPase and that in this study it was simply not detected, as even though BfRecN has the least sequence homology to EcRecN, it retains the amino acid residues essential to ATPase activity. Also, the conserved nature of RecN throughout the bacterial kingdom suggests an important functional role so why then would *B. fragilis* have a non-functional version of the protein in its genome, rather I simply failed to recreate the conditions in which it functions. AqRecN also failed to function in *E. coli*, even though it is a functional ATPase. However, since *A. aeolicus* is a hyperthermophile the protein may not fully function at 37°C, even though it does have ATPase activity at this temperature. Alternatively, as *A. aeolicus* is the most diverged species from *E. coli* studied (Deckert *et al.* 1998), the inability to function could represent an inability of AqRecN to interact with any putative EcRecN partners *in vivo*.



The ATPase activity of HiRecN and AqRecN appears to be DNA-independent but magnesium-dependent. Surprisingly, both the EcRecN and HiRecN carry a substitution of a Q-loop glutamine residue, which is supposed to be essential to  $Mg^{2+}$  binding (Hopfner *et al.* 2000). This was initially speculated to prevent ATP hydrolysis by preventing  $Mg^{2+}$  binding or alternatively that RecN possessed a novel ATPase mechanism quite distinct from that of the other ABC-type ATPases (Meddows 2002). However, the results clearly demonstrate ATPase activity that is dependent on  $Mg^{2+}$  ions. Perhaps the relatively conservative substitution of a histidine still allows  $Mg^{2+}$  binding, or RecN may possess an alternative  $Mg^{2+}$  binding site. Overall, the rate of ATP turnover by these two RecN proteins is comparable to that of the other SMC proteins, like cohesin, Rad50 and MukB (Hopfner *et al.* 2000; Hirano and Hirano 2004; Arumugam *et al.* 2006; Chen *et al.* 2008).

Both AqRecN and HiRecN were shown to cause linear ssDNA to be retained in the wells of polyacrylamide gels at high protein concentrations. This activity is dependent on the proteins having a functional ATPase activity. Although DNA association does not require a nucleotide to be bound, a mutation of the essential lysine residue in the Walker A box eliminated the DNA-RecN interaction. However, the observed DNA interaction is not DNA binding in a stochastic sense, as typically observed for known DNA binding proteins. The retention of DNA in the well suggests a large DNA-protein complex that is incapable of migrating into the gel. The data contrast with that reported for BsRecN. BsRecN binds 3' ssDNA and forms large, but discrete, nucleo-protein complexes in the presence of ATP (Kidane *et al.* 2004; Sanchez and Alonso 2005; Sanchez *et al.* 2006; Sanchez *et al.* 2007a). The formation of large nucleo-protein aggregates would explain the retention of material within the wells. Although this it is not a classical DNA binding activity, as it was blocked by addition of ADP, or mutation of the Walker A box, and was specific to ssDNA, it suggests that this the result of DNA-protein interactions. It is possible that the 'true' *in vivo* DNA target of RecN was not approximated and therefore DNA binding was not seen. However, it does seem likely that a region of ssDNA is required. Another possible, and very tempting suggestion,

is that the RecN proteins alone only weakly, or transiently, interact with DNA and that other factors, possibly partner proteins, are required to facilitate binding. This explanation is given more credence when considering that both Rad50 and BsRecN have been speculated to act as early sensors of DSBs, transiently associating and recruiting other factors to the break site.

Structural studies of RecN, based on sequence alignment and microscopy, may provide further insight into its activity and our understanding of how the protein functions. SMC proteins almost universally form dimers and this is essential to their function. Even the *B. subtilis* SMC, which can function without any other partners, forms a homodimer (Hirano and Hirano 1998b; Hirano 2005b). This leads to the speculation that RecN may also form dimers and that these could be the active form of the protein. Various evidence supports RecN protein forming complexes of some description. Firstly, during purification gel filtration suggested that the protein was much larger than the 60kDa expected for the monomer (Chapter 5). Secondly, electron microscopy showed what appear to be distinct AqRecN complexes, the structure of which is unclear, although, they differ from the classical V-shaped molecule commonly associated with SMC proteins (Melby *et al.* 1998; Haering *et al.* 2002). Experiments on BsRecN suggest that this protein also forms large complexes, which are associated with DNA (Sanchez and Alonso 2005; Sanchez *et al.* 2006; Sanchez *et al.* 2007a). The *E. coli* protein is also reported to have a propensity to form large protein aggregates *in vivo* and *in vitro* (Meddows 2002; Nagashima *et al.* 2006). During gel retardation assays, there was a small quantity of labelled material retained within the well, which would also suggest that the proteins are forming large non-migrating species *in vitro*. Finally, when the proteins were artificially cross-linked, several large species of RecN could be distinguished by SDS-PAGE analysis, including a dimeric species, supporting the idea that RecN may form complexes (Figure 6.13). However, it is not apparent whether RecN forms relatively simple dimers or large protein networks with many RecN proteins interacting.

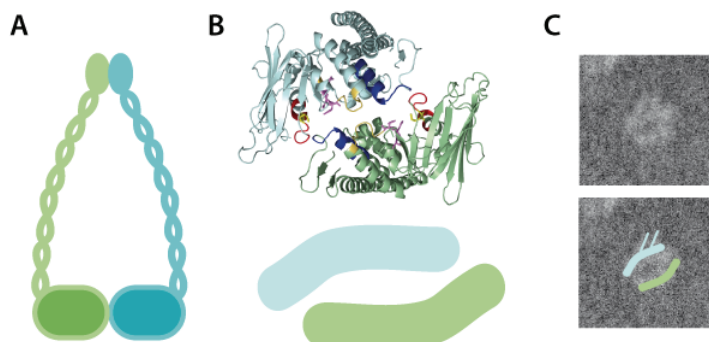
The RecN model developed by Dr. Geoff Briggs (University of Nottingham), which is based on two known SMC structures, illustrates the



possibility that dimerisation may be essential to function. As with many SMC proteins the key domains involved in ATP hydrolysis are not all in close proximity (Hopfner and Tainer 2003; Lammens *et al.* 2004). Specifically, the signature motif is located away from the Walker A and Walker B motifs. However, if the protein forms a dimer, the Signature motif of one RecN molecule could interact with the Walker motifs present on its counterpart and vice versa. This would form two functional ATP binding and hydrolysis pockets, and it can be clearly seen how the essential Walker A and Walker B motif residues involved in ATP binding and hydrolysis face into the pocket and could contact ATP (Figure 6.17). This nucleotide-sandwiched head domain structure would also potentially provide one of the interfaces required to dimerise the RecN proteins (Figure 6.17). In the future, the generation of mutations targeted to disrupt these potential dimer interfaces, and detect any phenotypic effects, could prove informative, especially since *in vitro* and *in vivo* assays to assess RecN activity are now available.

In all these cases, it is assumed that RecN is forming a dimer, via head and probably hinge domain interactions, like other SMC proteins (Figure 6.18 A). If this is true then it should be feasible to see this shape with the EM images of AqRecN. Although, these images are hard to interpret it is possible that what we are seeing are dimers. From the model structure it appears that the two RecN proteins pair and have a region that curves around slightly. If represented as a simple cartoon (Figure 6.18 B) it appears as a slightly curved sausage shaped structure. This shape, when overlaid onto the EM images does fit relatively well. The central hole of the doughnut shaped structure is actually the cavity, or at least a region of low density where the two RecN proteins meet and interface (Figure 6.18 C). There are two protrusions coming out of the top of this structure and facing away from it, which are assumed to be the coiled-coil region of the protein. If it is imagined that the RecN protein is pressed flat onto the grid then the coiled coils are also flattened, the globular domains form a circular structure as seen and the coiled-coil regions if laid down beside the globular domain form the protrusions. If they lie beneath or on top of the globular domain then they cannot be seen. Although this is just one

possible interpretation of the results seen it could explain how a RecN dimer gives rise to the structures seen.



**Figure 6.16 Diagram illustrating the potential for RecN to dimerise.**

The RecN molecules are shown as dimers in all cases one is light blue and one is light green.

- A)** Generalised model of an SMC protein dimer, potentially formed by RecN.
- B)** Model structure of RecN (Dr. G. Briggs) showing how dimerisation can facilitate formation of functional ATP binding pockets. Lower pane cartoon of the shape of this dimer.
- C)** EM image of AqRecN showing an example of one of the structure observed, the potential shape of a RecN dimer based on the model shown in B is overlaid onto the EM image.

In this chapter I have demonstrated that, like all SMCs, HiRecN and AqRecN are magnesium dependent ATPases. Furthermore, this activity is vital for their *in vivo* role in the repair of DSBs, although we cannot yet ascertain how RecN facilitates repair. It is interesting that some data contrasts with that reported for BsRecN, but it is possible that BsRecN has other cellular roles besides the repair of DSBs. For instance, *B. subtilis* is the only species in which RecN has been implicated in cellular competence (Kidane and Graumann 2005). The apparent structural and functional similarities between RecN and the eukaryotic SMC proteins are intriguing and of interest for future studies. It seems likely that RecN functions in a similar manner, forming dimers and potentially functional DNA-associating complexes. With the purification of

HiRecN, and the apparent ability of this protein to function within an *E. coli* cell, we have the tools to complement previous genetic studies. The value of this cannot be overstated and will allow future work to unravel the functions of RecN.

# Chapter 7

## General discussion

The central theme to this thesis has been the study of the repair of DNA breaks, and in particular the role of the RecN protein. Not long after the discovery of *recN* (Lloyd *et al.* 1983), evidence was presented to suggest that the RecN protein was involved in the repair of DSBs (Sargentini and Smith 1983; Picksley *et al.* 1984a). However, exactly what it did was not clear. Whatever its function, the fact that RecN is present in virtually all bacteria would suggest that it is of some importance.

My initial studies began with attempts to identify mutations that are synthetically lethal with *recN* and *recB* mutations. From these studies *ahpC* and *oxyR* mutations were demonstrated to be either synthetically sick or lethal in conjunction with mutation of several recombination genes (Chapter 3). Although H<sub>2</sub>O<sub>2</sub> is the least destructive ROS, it is likely to cause the majority of cellular damage due to its abundance (Storz and Imlay 1999; Seaver and Imlay 2001b; Seaver and Imlay 2001a; Seaver and Imlay 2004). The dependence of a cell upon recombination for survival in the absence of H<sub>2</sub>O<sub>2</sub> scavengers highlights the threat ROS pose and the importance of limiting DNA damage. However, as the majority of DNA damage caused by ROS is likely to affect just one DNA strand, it is unlikely that HR is required to repair DSBs caused directly by ROS, but rather to maintain replication forks and deal with replication dependent breaks that will arise. In such a background, where DNA damage is elevated due to the prevalence of ROS, it is conceivable that the increased levels of recombination can pose a threat to genomic stability. Even though recombination is an accurate pathway, errors can occur and increasing the potential for genomic rearrangements.

It was perhaps surprising that *recN* was not one of the mutations identified as being synthetically lethal with *ahpC* and *oxyR*. Although *recN* is not essential for recombination, in a wild-type genetic background, it can promote recombination as is evident from the recombination deficient

phenotype of a *recBC sbcBC recN* strain (Picksley *et al.* 1984a). RecN has also been suggested to increase the accuracy of repair by preventing break induced chromosomal rearrangements (Meddows *et al.* 2005). Therefore, it would be expected that the importance of RecN would be elevated in a genetic background where recombination is more prevalent. In fact, the screen for *recN* synthetic lethal mutants did not reveal any such mutation. This would suggest that such mutations don't exist (Chapter 3). However, the lack of synthetic lethal mutations does not rule out the possibility of *recN* interactions, or the existence of partner proteins *in vivo*, just that they are unlikely to be detected using this methodology.

The repair of a DSB is essential, as even a single break can be lethal. In bacteria, yeast and mammalian germ-line cells, where accurate repair is important for genome transmission, DSBs are corrected by HR. Largely due to studies in *E. coli*, the enzymology of this process is relatively well characterised. However, there are still proteins known to be involved, but with ambiguous roles in HR, RecN being one. RecN has been implicated in the repair of DSBs in a variety of species (Sargentini and Smith 1983; Picksley *et al.* 1984a; Alonso *et al.* 1993; Funayama *et al.* 1999; Kosa *et al.* 2004; Meddows *et al.* 2005; Wang and Maier 2008). Currently our understanding of the biochemistry of RecN, especially the *E. coli* protein on which the majority of the genetic studies were undertaken, is minimal. Therefore, a significant portion of this work is dedicated to the characterisation of the biochemical properties of RecN.

The *E. coli* RecN protein has proven to be of limited use for biochemistry, due to its insoluble nature (Meddows 2002; Nagashima *et al.* 2006). Data on the biochemistry of the RecN from *B. subtilis* is available, but the precise function of RecN remains unknown (Kidane *et al.* 2004; Sanchez and Alonso 2005; Sanchez *et al.* 2007a). As RecN is an ubiquitous bacterial protein, the possibility of studying a RecN ortholog, in place of the *E. coli* protein, was investigated (Chapter 4). Crucial to this approach was the identification of the capability of *H. influenzae recN* to encode a protein that could substitute *in vivo* for the *E. coli* protein. This *H. influenzae* protein, along

with two other RecN orthologs, from *A. aeolicus* and *B. fragilis*, was shown to be soluble after overexpression in *E. coli*. All three were purified to homogeneity (Chapter 5). Therefore, it was possible to characterise the biochemical activities of several RecN proteins for comparison, relate the activities to those already ascribed to BsRecN and, importantly, characterise a RecN protein that functions in *E. coli*, thus complementing previous genetic studies.

## 7.1 Suggested roles of RecN

The observation that RecN has structural similarity to the SMC proteins led to the speculation that it could possibly function in a manner comparable to at least one of the known SMC proteins. The studies of HiRecN and AqRecN would generally support this hypothesis. As expected, the RecN proteins are weak ATPases, with the ability to interact with DNA. They may also form complexes, although the nature of these has not been defined (Chapter 6). However, SMC proteins have diverse roles in DNA metabolism, many of which could be applied to RecN, although the key cellular roles of SMC proteins in eukaryotes appear to be provided in *E. coli* by another known SMC homolog, namely MukB (Niki *et al.* 1992; Graumann 2001; Sunako *et al.* 2001; Bartosik and Jagura-Burdzy 2005).

Eukaryotic cohesin is crucial for chromosome segregation, but also has roles in DNA repair, particularly the processing of DSBs and the maintenance of replication forks. The model of cohesin action is often used to explain the function of all SMC proteins. Certainly, the Smc5 – 6 complex is believed to function in a cohesin-like manner, trapping DNA within a ring like complex, and has been implicated in meiotic chromosome segregation, but appears to be primarily involved in DNA repair (Lehmann *et al.* 1995; Verkade *et al.* 1999; Pebernard *et al.* 2004). The third eukaryotic SMC complex is condensin. Although condensin has been implicated in DNA repair, specifically long-patch BER in vertebrates (Chen *et al.* 2004; Blank *et al.* 2006; Heale *et al.* 2006), it is primarily involved in chromosome compaction that occurs prior to segregation (Hirano 2005a). However, although I have shown in this study that

RecN proteins do share some functional homology with SMC proteins, this does not bring us closer to understanding its precise cellular function.

#### **7.1.1 RecN as a flag of DNA breaks**

It has been suggested from studies in *B. subtilis* that RecN protein binds a break soon after it occurs and acts as a flag to facilitate recruitment of other proteins (Kidane *et al.* 2004). This function is comparable to the role of Rad50, an SMC protein that forms part of the DSB sensing complex of eukaryotes (MR(N/X) complex) (Usui *et al.* 1998; Mirzoeva and Petrini 2001; Usui *et al.* 2001; McGowan and Russell 2004). As discussed in Chapter 1, this idea has some compelling support, as BsRecN forms nuclear-associated foci, even in the absence of end-processing (Sanchez *et al.* 2006). However BsRecN has been shown to only bind ssDNA, preferentially ssDNA that has an exposed 3' – OH group (Sanchez and Alonso 2005). This study failed to detect comparable ssDNA binding, but both HiRecN and AqRecN do interact with ssDNA. However, a break site will often have a blunt end with a single-stranded region for RecN binding only available after end-processing, by RecBCD. Secondly, RecN is only prevalent in a cell after SOS induction (Picksley *et al.* 1984b; Finch *et al.* 1985a; Rostas *et al.* 1987; Nagashima *et al.* 2006) and SOS induction only occurs after DNA damage, when ssDNA has accumulated (Higashitani *et al.* 1995). The small quantity of basally expressed protein would be adequate to act as a flag, which would explain the presence of only one or two foci, of RecN-fluorescent protein fusions on DNA during normal growth. So why is it then so heavily induced after damage? Moreover, why are multiple foci not then observed on the DNA (Moore; Kidane *et al.* 2004; Nagashima *et al.* 2006; Moore Unpublished data), as exposure to agents like IR or mitomycin C should cause a number of breaks in a dose dependent manner. It is possible that RecN acts as a flag of breaks, but it seems it must also have a second, later role in break repair that requires large quantities of the protein.

### 7.1.2 Potential of RecN to provide cohesion between DNA molecules

A second hypothesis is that the RecN functions in a cohesin-like manner. Cohesin holds sister chromatids together after replication and ensures correct segregation of the chromosomes. It has been speculated that cohesin, and the related Smc5 – 6 complex, aid the repair of a DSB by holding sister chromatids in proximity. This would increase the chance of homologous sequences being in proximity, thus promoting efficient and accurate repair by HR (Mengiste *et al.* 1999; Skibbens *et al.* 1999; Sjogren and Nasmyth 2001; Morishita *et al.* 2002; Lehmann 2005; Strom and Sjogren 2007). If RecN were a cohesin-like molecule then it would be expected to topologically interact with DNA. This could explain the limited DNA binding seen with HiRecN and AqRecN as cohesin is believed to entrap and encircle rather than bind DNA (Haering *et al.* 2002). However, cohesin requires ATP hydrolysis to allow its loading onto DNA (Arumugam *et al.* 2003) and to facilitate complex formation, possibly by stabilising the dimer interface within the head domain of SMC proteins (Weitzer *et al.* 2003; Lammens *et al.* 2004). In contrast, the two RecN proteins analysed in this study, associate with DNA in the absence of a nucleotide, with the interaction being inhibited in the presence of ADP. Similarly, BsRecN could also interact with DNA, in the absence of a nucleotide (Sanchez and Alonso 2005; Sanchez *et al.* 2007a). Perhaps, as suggested in Chapter 6, RecN requires ATP hydrolysis to allow it to release DNA. Therefore, in the presence of ADP all of the RecN is in a conformation that prevents DNA interaction and consequently cannot load onto the DNA. Interestingly, the *E. coli* SMC homolog, MukB, and BsSMC have been shown to bind DNA in an ATP-independent manner (Niki *et al.* 1992; Hirano *et al.* 2001). Therefore, ATP-independent DNA binding would not be unique within the SMC family of proteins to RecN, instead SMC proteins appear to have diverse DNA interactions and requirements.

As ATP binding and hydrolysis is not required by either SMC proteins, or more particularly RecN to bind DNA it therefore seems contradictory that the K35 mutants generated did not bind DNA. Although it could be that the



mutant proteins fail to fold correctly, this seems unlikely. Instead I believe the explanation lies in the disruption of a potential dimer interface. If RecN is functioning in a manner related to cohesin disruption of the dimer would prevent it interacting with and trapping DNA. The simplest explanation for the deficiencies observed for the K35A mutations is that if RecN does function as a dimer, with at least one interface formed by the interaction of the ATPase domains (Chapter 6), that the ATPase mutations may disrupt this interface. The disruption may prevent dimerisation and thus the function of RecN. Such disruption has been observed with several SMC proteins (Lammens *et al.* 2004).

What is curious is the lack of dsDNA binding observed for RecN proteins. If the protein engages DNA by entrapment, then it should be expected to bind dsDNA in a comparable manner to its binding of ssDNA. However, neither I nor the authors of the reports on BsRecN detected any such activity (Sanchez and Alonso 2005; Sanchez *et al.* 2007a). Evidence suggests that BsSMC does preferentially bind ssDNA over dsDNA (Hirano and Hirano 1998a; Sanchez and Alonso 2005). However, BsSMC still binds to dsDNA and the mechanism by which BsSMC preferentially binds ssDNA is unknown. But perhaps RecN binds to ssDNA preferentially or exclusively using a similar mechanism. However, further studies are needed to define the DNA substrates to which RecN can bind and exclude the possibility of a dsDNA binding activity.

### **7.1.3 RecN as a condensin**

The putative models for RecN activity are largely based on the proposed mechanisms of cohesin activity. However compelling the cohesin DNA entrapment model is, it is not the only model and issues remain unresolved (Hirano 2002; Huang *et al.* 2005; Guacci 2007). The eukaryotic condensin complex is believed to function in a distinct manner from cohesin. DNA entrapment appears to be unlikely, as disruption of the proposed condensin ring does not affect the DNA binding activity of the complex (Hudson *et al.* 2008). Rather, it is believed that condensin shares functional homology with the *E. coli* SMC homolog, MukB. This protein also forms a complex, with Muke

and MukF, and this MukBEF complex appears to fulfill the role of condensin in *E. coli* (Ohsumi *et al.* 2001). It is proposed that ATP-bound MukBEF molecules polymerise along DNA in a cooperative fashion to form a filament that compacts the bound DNA molecule. This is achieved via interactions between the heads of adjacent MukB molecules along the DNA, causing it to coil (Case *et al.*, 2004), and this is dependent on ATP hydrolysis (Chen *et al.* 2008). Quite how a compaction of the chromosome would aid recombination is unclear and a cohesion-like mechanism of RecN action, with the tethering of DNA molecules together, seems the more likely model. However, the possibility that RecN may bind DNA in a condensin like manner and polymerise on DNA cannot be ruled out. A requirement to polymerise would explain the tendency of RecN to aggregate and form complexes (Kidane *et al.* 2004; Nagashima *et al.* 2006). However, in terms of the role it fulfills, i.e. the repair of DNA, particularly DSBs, it most closely resembles the cohesin and Smc5 – 6 complexes, both of which appear to act by entrapment of DNA.

## **7.2 RecN and SMC proteins can aid accurate genome transmission**

To ensure the proper transmission of its genome, a cell must correctly replicate, segregate and, if necessary, repair its DNA. The cohesin and Smc5 – 6 complexes have key roles in all of these processes. Mutations in components of the cohesin complex have been shown to be mildly sensitive to UV and critically sensitive to IR exposure (Birkenbihl and Subramani 1992). This is similar to *E. coli* cells carrying a *recN* mutation, which show a very mild UV sensitisation, but are markedly IR sensitive (Figure 4.7) (Sargentini and Smith 1983; Picksley *et al.* 1984a). The IR sensitivity of cohesin suggests a role in the repair of DSBs. This has been confirmed in a variety of organisms.

Sjörger and Nasmyth (2001) showed that in yeast cells, cohesin has to be present and active, otherwise cells cannot repair IR induced DSBs. Activation requires DNA replication and the Eco1 protein. In the absence of Eco1, cells show an IR sensitive phenotype that is almost identical to that reported by Sjörger and Nasmyth (2001) for cohesin deficient cells (Skibbens

*et al.* 1999; Toth *et al.* 1999; Sjogren and Nasmyth 2001). In chicken DT-40 cells depletion of the cohesin component Scc1, causes sensitivity to IR. IR exposure also drastically increased the incidence of chromosomal rearrangements, suggesting that although repair occurs, it was inaccurate. It appears that in the absence of sister chromatid cohesion, genomic stability is seriously compromised (Sonoda *et al.* 2001). Direct evidence for cohesin having a role in the repair DSBs comes from a very elegant set of experiments, by Kim *et al.* (2002). They used human (HeLa) cells and a laser beam, at a specific frequency that will only damage DNA. The beam caused tracks of DSBs through the targeted cell. Using fluorescently tagged proteins, they monitored the recruitment of proteins to the break tracks and saw that cohesin complex components were recruited to these tracks of DSBs, in an Mre11 dependent manner. Mre11 is part of the MR(N/X) complex, which is known to detect and aid processing of a DSB and to which RecN has already been compared, as an early sensor or flag of DNA breaks (Sanchez *et al.* 2006). The observations by Kim *et al.* (2002) suggest that cohesin is actively recruited to, and required for the repair of a DSB and can be detected in a region of approximately 50-100kb around the break site (Kim *et al.* 2002a; Strom *et al.* 2004; Unal *et al.* 2004).

The most recently identified SMC containing complex, the Smc5 – 6 complex, is also heavily involved in DNA repair. The Smc5 and Smc6 proteins were identified in *S. pombe*, as mutants sensitive to UV and IR (Nasim and Smith 1975; Phipps *et al.* 1985; Fousteri and Lehmann 2000). These mutants were also shown to be compromised for NER and DSB repair (Lehmann *et al.* 1995; Verkade *et al.* 1999). These SMC proteins were also shown to be part of a larger complex, with at least six non-SMC subunits (Nse1 through Nse6). Interestingly two of the subunits, Nse5 and Nse6, appear to form a separate sub-complex, unusually this does not interact with the head domains, but binds within the coiled-coil region of the SMC proteins. Also, Nse2 is a SUMO ligase, directly involved in telomere maintenance and while in *S. cerevisiae* all six Nse proteins are essential, in *S. pombe* the sub-complex forming Nse5 and Nse6 are dispensable (Fousteri and Lehmann 2000; Andrews *et al.* 2005; Potts

and Yu 2005; Sergeant *et al.* 2005; Zhao and Blobel 2005; Palecek *et al.* 2006; Potts and Yu 2007).

The Smc5 – 6 complex has been implicated specifically in the repair of DSBs, although is involved in HR in general (Mengiste *et al.* 1999; Lehmann 2005; Strom and Sjogren 2007). In cells deficient for components of the complex, DNA fragmentation occurs. This was presumed to be due to the accumulation of DSBs and Cost and Cozzarelli (2006) substantiated this claim by experiments that showed the fragmentation to be exacerbated by exposure to IR (Cost and Cozzarelli 2006). This is surprisingly reminiscent of what is seen with *E. coli recN* mutants. After IR exposure fragmentation of the *E. coli* chromosome normally occurs, however, it is more severe in cells carrying a *recN* mutation (Morishita *et al.* 2002). Surprisingly, while cohesin is required to initiate recombination, the Smc5 – 6 complex is essential to the resolution of recombination, specifically the crossover structures that occur within regions of repetitive sequence (called rDNA regions). This could explain the delay in chromosome segregation observed in Smc5 – 6 deficient cells (Torres-Rosell *et al.* 2005; Lindroos *et al.* 2006), while cohesin mutants segregate prematurely. The inability to segregate would also explain the requirement for the Smc5 – 6 complex during meiotic chromosome segregation, which requires both cohesin and the Smc5 – 6 complex. The Smc5 – 6 complex is not acting as a cohesin in this case, but is needed to allow resolution of crossovers (Pebernard *et al.* 2004). Perhaps the most persuasive evidence that the Smc5 – 6 complex is involved in the repair of DSBs is that it is actively recruited to them, in an Mre11-dependent manner, much like cohesin (De Piccoli *et al.* 2006; Lindroos *et al.* 2006; Potts *et al.* 2006).

Both cohesin and the Smc5 – 6 complex appears to share some overlapping functions, and currently share the same model of action. Both are required for accurate segregation of chromosomes and the distribution of the Smc5 – 6 complex is altered by disruption of cohesin loading (Lindroos *et al.* 2006). However, since both complexes are essential it would be expected that they must possess distinct roles in DNA metabolism. Perhaps the cohesin provides cohesion on a genome-wide scale, although it is recruited to break

sites to supplement cohesion in this region, whilst the Smc5 – 6 complex has a dual function, providing temporary cohesion at the break site and a role in resolving crossovers, facilitating HR and repair of a DSB.

The simplest explanation for cohesin's role would be that the complex holds sister chromatids together around the breaks site and this could aid HR by keeping homologous regions in close proximity (Sjogren and Nasmyth 2001). An early role in HR would explain why genetic studies failed to place *recN* in either the RecFOR or RecBCD recombination pathway. It was placed in its own epistatic grouping, as it acts before either pathway is employed (Lloyd and Buckman 1991). However, unlike cohesin mutants, recombination still proceeds in the absence of RecN. Therefore, RecN is not essential to the initiation of recombination. Yet, this may be a result of the complexity, and size, of eukaryotic genomes making cohesion a requirement to ensure accurate repair, rather than a difference in role between cohesin and RecN.

It is eminently plausible that RecN, like cohesin and the Smc5 – 6 complex, is recruited to a DSB, and like cohesin this occurs early during repair. Evidence suggests that BsRecN may in fact be the first protein recruited to a break, forming the basis of the proposed 'repair centres' (Kidane *et al.* 2004). However, why only one or two RecN foci are observed on the DNA, even after chronic DNA damage will have induced numerous breaks, is unclear (Kidane *et al.* 2004; Nagashima *et al.* 2006; Moore Unpublished data). Perhaps the suggestion by Kidane *et al.* (2004) of repair centres, with numerous breaks being processed in each, is the answer. This hypothesis seems unlikely, as when numerous recombination events occur in close proximity, the potential for errors should be magnified. However, the results are complicated by the inability of the EcRecN fluorescent proteins to fully complement for a deletion of the *recN* gene (Moore Unpublished data) and the tagged proteins are expressed from a plasmid. Future work may make use of chromosomal based fusions, regulated by the *recN* promoter, to try and overcome this problem.

### **7.3 RecN and SMC proteins are involved in DNA replication**

RecN is indirectly implicated in the maintenance of replication, as this is believed to be the housekeeping function of HR (Cox *et al.* 2000; Mariani 2000; Sandler and Mariani 2000; Mariani 2004). However, it has been more directly implicated as it is required to facilitate break induced replication of a non-replicating plasmid (Asai *et al.* 1994).

The Smc5 – 6 is known to be involved in DNA replication. As depletion of the complex causes sensitivity to hydroxy urea, this could be suppressed by deletion of Rad51. This would suggest that without the Smc5 – 6 complex, collapsed replication forks either undergo illegitimate recombination, or more likely, HR is initiated, but cannot be resolved in the absence of the complex (Harvey *et al.* 2004; Ampatzidou *et al.* 2006). It also appears that loading of the Smc5 – 6 complex onto DNA is associated with passage of the replication fork (Tsuyama *et al.* 2006). Similarly, cohesin mutants show sensitivity to hydroxy urea (Tatebayashi *et al.* 1998) and the establishment of sister chromatid cohesion requires the passage of a replication fork. It appears that cohesin is normally loaded onto chromosomes at specific locations, referred to as cohesin-associated regions (CARs), prior to DNA replication (Blat and Kleckner 1999; Glynn *et al.* 2004; Lengronne *et al.* 2004). As DNA replication occurs, cohesin is activated enabling sister chromatid cohesion and the activation requires the Eco1 (sometimes called Ctf7) protein. How Eco1 activates sister chromatid cohesion is unclear, but there is evidence that it may interact directly with components of the replisome. It also appears that Eco1 mediated activation is essential, as Eco1 deficient cells have phenotypes almost identical to cells deficient for cohesin (Skibbens *et al.* 1999; Toth *et al.* 1999; Lengronne *et al.* 2006; Moldovan *et al.* 2006). It is believed that cohesin-associated regions provide spatial localisation, by specifying where cohesion binds, while Eco1 activity allows temporal activation of sister chromatid cohesion, after DNA replication has occurred.

RecN shares a remarkable functional resemblance to the cohesin and Smc5 – 6 complexes, possessing a key role in the repair of DNA DSBs, and

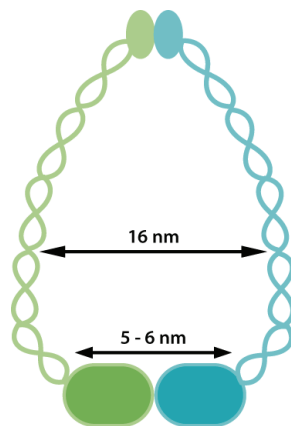
more generally in HR. However, unlike the eukaryotic SMC complexes, RecN is not essential and has no known interaction or complex forming partners. It is also interesting that the unique V-shaped structure associated with SMC dimers (Melby *et al.* 1998) could not be visualised under an electron microscope. This may be because RecN is much smaller and so less distinct, or that it forms different shaped complexes. The formation of large complexes would explain the protein aggregates observed in *E. coli* and the large nucleoprotein aggregates BsRecN forms with ssDNA (Meddows 2002; Sanchez and Alonso 2005; Nagashima *et al.* 2006; Sanchez *et al.* 2007a). In future, the search for interaction partners via “pull down” experiments could be informative. It seems likely that RecN functions in a manner comparable to one or more of the SMC complexes. However, its cellular role maybe distinct from that of those currently characterised for SMC proteins.

## 7.4 Models of RecN function

Any model for RecN that suggests a cohesin-like function has to account for the ssDNA binding activity of the protein, the essential ATPase activity and how cohesion of DNA molecules by RecN could aid the rapid and accurate repair of a DSB.

For the models I am about to propose, several assumptions are being made. Firstly, RecN forms a dimer that allows the formation of functional ATPase domains (Chapter 6) and also a ring shaped structure (Figure 7.1). Whether RecN requires partner proteins to form such a complex is speculative, and it maybe that, like BsSMC, it can act alone as a simple homodimer (Hirano and Hirano 1998a). Secondly, that the tendency of the protein to aggregate and form large complexes is physiologically relevant. Thirdly, that RecN not only has a cellular role like cohesin, but also acts in a similar manner, functioning by entrapment of DNA within the ring of the homodimer (Haering and Nasmyth 2003). Finally, SOS induction of RecN is required to provide enough RecN to allow cohesion to occur, potentially on a genome-wide, and certainly on a break-localised scale.

#### 7.4.1 Break localised cohesion model of RecN activity



**Figure 7. 1. Generalised structure of a RecN dimer.**

The two RecN molecules are shown, with dimerisation occurring at the head and hinge region. The ring structure is highlighted with the estimated diameter shown.

Kidane *et al.* (2004) suggested that RecN could detect a break and promote repair by binding to it and flagging the break, so as to recruit repair factors. RecN may potentially then act in a cohesion-like manner and trap the intact homolog, thus ensuring that the repair template is in close proximity to the break, promoting the speed and accuracy of repair. The basal expression of RecN ensures there are enough molecules present in the cell to act as a flag, binding a break and then directing repair to this site (Figure 7.2 A), by recruiting other repair proteins. The ‘flag RecN dimer’ could also act as a nucleation centre so that further RecN dimers can polymerise or aggregate with this RecN molecule, localising RecN to the break site (Figure 7.2 B). Although the original flagging and nucleation events would occur prior to DNA end-processing, the subsequent action of RecN and the localised break-induced cohesion could occur after end-processing, when ssDNA would be available.

RecN has been shown to bind or interact with ssDNA by Sanchez and Alonso (2005) for BsRecN and as demonstrated with HiRecN and AqRecN in this study. It can be envisaged that RecN, localised to the break site by nucleation with the ‘flag RecN’, might then specifically interact with ssDNA in this region. It is presumed that RecN will interact, and thus capture the processed 3' ssDNA tail at the break site, facilitating transports of the ssDNA tail into the interior of the RecN dimer ring (Figure 7.2 C). The relatively weak ssDNA interactions of RecN proteins seen in this work (Chapter 6) could be



due the small linear substrate used. If RecN binding to ssDNA is to facilitate ssDNA capture, rather than a direct binding activity, then the linear substrate can be held weakly but once within the RecN ring it will not be held, therefore the substrate would be free to migrate. It is only due to the excessive quantities of protein used that we see the retention of labelled material within the well in gel retardation assays, as a result of the initial ssDNA capture by RecN. The formation of a large RecN polymer would also explain why the labelled material remained in the well.

Once the RecN polymer has captured the ssDNA tail, it might then also capture, in a non-specific manner, any DNA molecules nearby, transporting them into the interior of the RecN dimer ring (Figure 7.2 D). The capture of a sister chromatid assumes that it has not yet had time to move and will still be nearby. If this is the case then the recombinogenic ssDNA tail and the intact sister, which can act as a template for repair, will be held in proximity. The cohesin ring has been estimated to have a diameter of approximately 50nm, RecN would be much smaller as its coiled-coil domains are only about a third the size, although based on the size and circumference of a cohesin ring, a RecN dimer ring could have a diameter of up to 16nm. Although direct measurement from the model of the RecN dimer suggests a gap of only 5-6nm, which is much smaller than was estimated from the size of the cohesin ring. However, the coiled-coil motifs are flexible and so can move outward to give the larger radius away from the head. As a RecA filament is around 10nm (Yu and Egelman 1992), it could be possible for RecA to begin loading onto the ssDNA without too much interference from the RecN ring and initiate strand invasion. It is also possible that although RecN polymers are bound to the ssDNA tail they cover only a fraction of it, leaving the majority of the tail free to initiate recombination. Whether RecN remains bound throughout recombination is unclear. The ring's diameter would be adequate to contain a RecA filament and duplex DNA, but its presence may still interfere with recombination and it is possible that RecN will release the DNA, or be degraded by ClpXP, at this point allowing recombination to proceed unhindered (Figure 7.2 E).

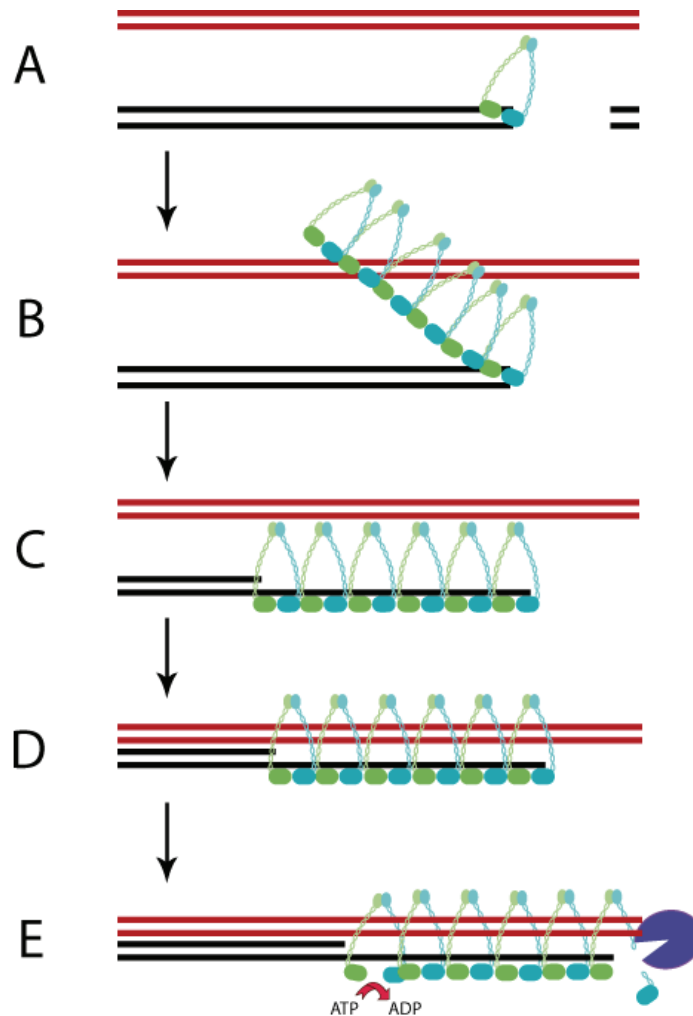


Figure 7. 2. Model for the involvement in RecN in providing break recognition and break specific cohesion.

Two DNA duplexes are shown with the break in one (black) and the intact sister (red).

- A)** RecN dimer (light blue and green dimer) can recognise and bind the DNA end acting to 'flag' the break.
- B)** The flag RecN remains bound at the end and acts as a nucleation point for a RecN polymer to form.
- C)** End-processing exposes the 3' ssDNA tail, which can be bound by RecN dimers resulting in capture of the tail.
- D)** The RecN dimers then capture the intact duplex bringing it into the RecN ring, holding the intact sister in proximity to the broken duplex.
- E)** RecN is removed from the DNA, either by proteolysis by ClpXP (blue circle) or the ATPase activity of RecN disrupting the dimer, to allow recombination to proceed. Strand invasion could commence prior to clearance.

A release hypothesis could explain the requirement for ATPase activity. These studies suggest that RecN can interact with ssDNA with or without ATP, but ADP blocks this interaction. The explanation was that RecN, in the presence of ADP, is held in a conformation that promotes release of DNA (Chapter 6). If this is the case then RecN could bind ssDNA and trap nearby duplex without ATP hydrolysis. Turnover of ATP to ADP could then drive the conformation change that converts RecN into a non-DNA binding conformation and thus it will release the DNA, allowing recombination to proceed. The relatively weak ATPase activity could therefore be due to a need to hold DNA for a certain time period and then release. It could also be that other factors can regulate the ATPase activity to control when RecN disassociates from the DNA. A release mechanism could also prevent the basal levels of RecN disrupting replisome progression or access to the DNA, although such disruption would only seem likely when large aggregates are present. Once released the RecN could then be degraded by ClpXP. This scenario would explain why the ATPase deficient K35A mutations of RecN failed to function *in vivo* (Chapter 6), as ATP hydrolysis is needed to release DNA, without which RecN cannot function correctly. However, the above scenario does not explain why the mutant RecN proteins fail to bind DNA, but the simplest explanation would be the disruption of the dimer interface, as suggested earlier (Section 7.1.2).

#### **7.4.2 Global cohesion model of RecN activity**

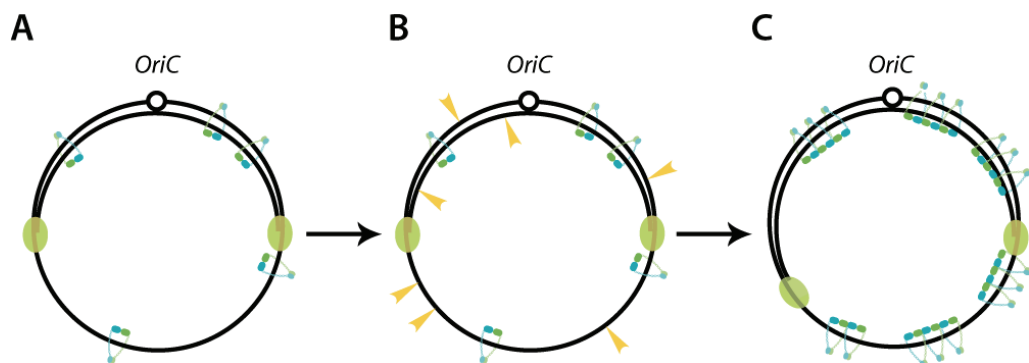
In a modification to the specific cohesion model of RecN action, it is also possible that RecN could provide genome-wide cohesion in response to DNA damage. This model was originally proposed by Meddows (2002) to explain why RecN is so heavily induced during the SOS response and the idea has been given credence by recent reports that DNA damage can induce sister chromatid cohesion in yeast outside of anaphase (Strom and Sjogren 2005). In eukaryotes, damage-induced sister chromatid cohesion occurs on a global, genome-wide scale with a single DSB inducing cohesion not just on the chromosome where the break occurs, but across all chromosomes. The activation of this genome-wide cohesion is dependent on Eco1, but does not

require DNA replication (Strom *et al.* 2007; Unal *et al.* 2007). It appears that, although, cohesin normally provides a general sister chromatid cohesion, temporally activated by replication and spatially organised by CARs, after DNA damage this cohesion is enhanced on a genome-wide scale in order to aid repair, particularly of DSBs. Whether this global cohesion is random, or remains spatially organised is unclear, although the active recruitment of cohesin to the site of a break suggests at least some specificity (Kim *et al.* 2002a).

In *E. coli* the newly replicated chromosome are normally held together immediately after replication by the catenation of the DNA molecules, which occurs during replication. This provides cohesion at a particular locus for approximately 15 minutes after its replication. The catenation is then removed by the action of topoisomerase and the chromosomes can begin to segregate (Wang *et al.* 2008). This ensures that sister chromosomes are held in proximity immediately post-replication. However, if damage occurs then segregation would separate the sister chromosomes making repair by HR difficult. Therefore, there is a need in the cell for a mechanism to keep sister chromosomes together, particularly after damage to facilitate repair by HR. If RecN is a cohesin-like molecule then it might do this, and like cohesin provide genome-wide cohesion in response to DNA damage. By holding together the newly replicated chromosomes, RecN could facilitate the search for homology and thus the accuracy and rate of repair.

In this global cohesion model, the basally expressed RecN is loaded onto the chromosome. This could be at random or, as with cohesin, targeted to specific regions of the chromosome. In either case RecN is distributed around the chromosome. The presence of individual RecN dimers encircling the DNA will not affect replisome progression, either they get 'pushed' along in front, or perhaps the ATPase activity allows them to release the DNA when a replisome encounters them. However, the randomly loaded RecN dimers can act as nucleation centres for the assembly of RecN aggregates around the chromosome. These numerous nucleation sites could then provide spatial distribution of RecN around the chromosome. There is also the possibility that

RecN has specific loading sites, like the CARs associated with cohesin (Blat and Kleckner 1999; Glynn *et al.* 2004; Lengronne *et al.* 2004). Either mechanism will distribute RecN dimers around the chromosome (Figure 7.3 A). In response to DNA damage (Figure 7.3 B), RecN expression is induced as part of the SOS response. The newly induced RecN can then associate with the preloaded RecN and assemble a polymer of RecN, which will capture nearby DNA molecules and thus allow global cohesion between the parental and newly replicated genomes (Figure 7.3 C).



**Figure 7. 3. Model for the involvement of RecN in global chromosome cohesion.**

- A) Basally expressed RecN is loaded onto the DNA duplex (black lines) as replication proceeds (replisome, green oval) but does not block replication.
- B) DNA damage (yellow arrows) occurs and the SOS response is induced.
- C) RecN is induced and the pre-loaded RecN acts as nucleation centres for RecN aggregates/polymers that then act to hold the parental and newly replicated DNA in proximity. The polymers of RecN can block or slow replisome progression until repair is complete.

It is also possible, although highly speculative, that these RecN polymers could inhibit replisome progression, stalling replication until repair is complete and reducing the risk of creating further DSBs or fixing mutations. If we assume that RecN entraps DNA molecules then it will be ‘wrapped around’ the DNA. It would seem unlikely that RecN could limit access to the DNA. Even though estimates from the expression data on RecN suggest there could be 10,000 copies per cell after SOS induction (Finch *et al.* 1985a). The assumption is that RecN is a dimer and if randomly distributed then there would only be one RecN dimer every kilobase. It is also known from experiments using fluorescently tagged proteins that much of the RecN appears

to be in the cytoplasm as aggregates (Nagashima *et al.* 2006; Moore Unpublished data). However, if RecN does polymerise or aggregate on the DNA then it may block access to the DNA beneath the polymer and more importantly could restrict the advancement of a replication fork. An array of RecN dimers could be generated as a result of the polymerisation of RecN or via the gathering of RecN rings in front of the replisome as it travels, accumulating more rings as it proceeds. This array could eventually stall replication and potentially collapse the replication fork. In this way RecN could act as a checkpoint, preventing DNA replication and cell division until repair has been completed and the RecN protein removed.

Inhibition of replication by RecN would explain the stringent regulation of, and transient expression of, RecN. At least two known, and one putative, LexA binding sites normally repress expression of RecN (Rostas *et al.* 1987; Erill *et al.* 2003). Once DNA damage occurs repression is removed and *recN* gene is rapidly expressed, peaking within 20 minutes (Nagashima *et al.* 2006). The presence of this RecN could allow cohesion to facilitate HR and also act to prevent replication. This could help prevent the risk of either passing on errors or complicating the repair of any breaks. Within 160 minutes of SOS induction RecN is completely removed, largely by the action of the ClpXP protease (Nagashima *et al.* 2006; Neher *et al.* 2006). This would free the replisomes to proceed and cell division to resume. If RecN does act as a replication block it provides an explanation for the toxicity observed when RecN is overexpressed from a multi-copy plasmid (Meddows 2002). Cohesin could have a similar affect and this would explain the stringent regulation of cohesins' activation, specifically after DNA replication. Therefore, it could be assumed that prior to DNA replication the presence of active cohesin would be detrimental to the cell (Toth *et al.* 1999; Lengronne *et al.* 2006). If RecN only exists as an active form, it must be tightly regulated and expression controlled, to avoid a toxic affect on the cell and to ensure its removal it is actively targeted by a protease (Nagashima *et al.* 2006; Neher *et al.* 2006).

Although two models of RecN action during DSB repair have been suggested; it is possible that what actually occurs within the cell is a hybrid of

both models, with RecN providing break-localised and genome-wide cohesion in response to DNA damage, much like the damage induced cohesion associated with cohesin (Strom and Sjogren 2005; Strom *et al.* 2007; Unal *et al.* 2007). The important point is that if RecN dimers can trap DNA molecules, they will keep the trapped molecules in close proximity to each other, facilitating repair. This will ensure that homologous sequences are kept in close proximity and thus the template for the repair of a break will be nearby. This will increase the ease and speed of repair, as well as the accuracy. Such a theory nicely explains how RecN is involved in the repair of DSBs, but also why the protein is not essential. In the absence of RecN recombination can still occur, as RecA will still find and invade homologous sequences. But the model predicts that the time taken to find the homologous sequences, and thus the chance of making errors, is increased.

## 7.5 Future work

Although the evidence presented in this thesis has provided further insight into the RecN protein and its function, the mechanics of its activity remains unresolved. A model for RecN activity is suggested in which RecN functionality is related to the cohesin model of DNA entrapment, however, it is by no means certain that RecN has a cohesin-like role. The same model is also used to explain the activity for the related, Smc5 – 6 complex, which has distinct cellular roles different from cohesins. With the isolation and purification of HiRecN it is possible for the first time to compare the biochemical activity of RecN to the extensive genetics performed in *E. coli*. I have also begun the process of characterising the activities of RecN, showing that it is a weak, magnesium-dependent ATPase that can interact with ssDNA. This has led to a closer comparison of RecN to other SMC family proteins and it is now possible to more clearly define RecN as an SMC protein, due to the shared activities. However, work remains to be done. RecN is an ubiquitous bacterial gene that whilst not essential, is clearly important when a cell is stressed. While a model for RecN activity is presented here, work is needed to test the ideas presented. For instance, there is no explanation for why multiple

RecN foci are not seen on the DNA, either as a result of global cohesion or at site of multiple breaks (Kidane *et al.* 2004; Nagashima *et al.* 2006; Moore Unpublished data).

One approach is to attempt to localise where RecN acts and determine if it is loaded onto the DNA prior to damage. It is possible to detect chromosomally associated proteins using chromosome immune-precipitation (ChIP) assays. To this end, the purification of highly specific RecN antibodies would be useful. During this work, I raised antibodies to the insoluble EcRecN, but these showed limited specificity to RecN and were therefore inadequate to allow quantification of the protein or further analyse its location and function. However, antibodies raised to the N-terminal globular domain have shown more promise (Grove, unpublished). Antibodies may also allow pull down assays of RecN to try and identify any interaction partners. These partners could be part of a RecN ring complex like those associated with the eukaryotic SMC proteins, or may act to regulate the ATPase activity of RecN, like the ScpA and ScpB proteins that interact with BsSMC, perhaps to promote release of DNA (Schleiffer *et al.* 2003; Hirano and Hirano 2004; Palecek *et al.* 2006; Petrushenko *et al.* 2006b). If partner proteins do exist and they may have other cellular roles, like the Nse subunits of the Smc5 – 6 complex and therefore would not be readily detectable by screening for recombination deficiencies related to RecN as the deficiency may get masked by loss of alternative functions. It is also possible to make use of functional His-tagged HiRecN to do pull downs directly from *E. coli* cell extracts as HiRecN has been demonstrated to fully complement for the EcRecN *in vivo* it should interact with any *E. coli* partner proteins (Chapter 4). This provides an alternative method to isolate any RecN partners.

Another method to discover if RecN does form complexes would be to try and investigate its structure. The EM studies provide evidence for the formation of higher order AqRecN structures and glutaraldehyde cross-linking and gel filtration data suggests that the RecN proteins do form complexes with themselves, but the exact composition is still unknown (Chapter 5; Chapter 6; and Kidane *et al.* 2004). Work is also progressing on crystallisation trials and



early screening using AqRecN and HiRecN has been promising. However, further work is needed to isolate ideal conditions and crystallise the protein. The resolution of the crystal structure is particularly valuable, as there is currently no crystal structure for a full-length SMC protein (van den Ent *et al.* 1999; Lowe *et al.* 2001; Lammens *et al.* 2004; Fennell-Fezzie *et al.* 2005). Therefore, a crystal structure of the full-length RecN protein could aid our understanding not only of RecN, but the whole SMC protein family.

Although how RecN aids the repair of a DNA break remains unknown, this study has provided new information and importantly, the tools and raw material needed to begin fully characterising and understanding RecN, work which is continuing in the Lloyd laboratory.

# Chapter 8

## References

- Allers, T. and Lichten, M. 2001a. Differential timing and control of noncrossover and crossover recombination during meiosis. *Cell* **106**(1): 47-57.
- . 2001b. Intermediates of yeast meiotic recombination contain heteroduplex DNA. *Mol Cell* **8**(1): 225-231.
- Alonso, J.C., Stiege, A.C., and Luder, G. 1993. Genetic recombination in *Bacillus subtilis* 168: effect of *recN*, *recF*, *recH* and *addAB* mutations on DNA repair and recombination. *Mol Gen Genet* **239**(1-2): 129-136.
- Ampatzidou, E., Irmisch, A., O'Connell, M.J., and Murray, J.M. 2006. Smc5/6 is required for repair at collapsed replication forks. *Mol Cell Biol* **26**(24): 9387-9401.
- Anderson, D.E., Losada, A., Erickson, H.P., and Hirano, T. 2002. Condensin and cohesin display different arm conformations with characteristic hinge angles. *J Cell Biol* **156**(3): 419-424.
- Anderson, D.G. and Kowalczykowski, S.C. 1997. The translocating RecBCD enzyme stimulates recombination by directing RecA protein onto ssDNA in a chi-regulated manner. *Cell* **90**(1): 77-86.
- Andrews, E.A., Palecek, J., Sergeant, J., Taylor, E., Lehmann, A.R., and Watts, F.Z. 2005. Nse2, a component of the Smc5-6 complex, is a SUMO ligase required for the response to DNA damage. *Mol Cell Biol* **25**(1): 185-196.
- Aravind, L., Walker, D.R., and Koonin, E.V. 1999. Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res* **27**(5): 1223-1242.
- Arumugam, P., Gruber, S., Tanaka, K., Haering, C.H., Mechtler, K., and Nasmyth, K. 2003. ATP hydrolysis is required for cohesin's association with chromosomes. *Curr Biol* **13**(22): 1941-1953.
- Arumugam, P., Nishino, T., Haering, C.H., Gruber, S., and Nasmyth, K. 2006. Cohesin's ATPase activity is stimulated by the C-terminal Winged-Helix domain of its kleisin subunit. *Curr Biol* **16**(20): 1998-2008.
- Aruoma, O.I., Kaur, H., and Halliwell, B. 1991. Oxygen free radicals and human diseases. *J R Soc Health* **111**(5): 172-177.
- Asad, L.M., Asad, N.R., Silva, A.B., de Almeida, C.E., and Leitao, A.C. 1997. Role of SOS and OxyR systems in the repair of *Escherichia coli* submitted to hydrogen peroxide under low iron conditions. *Biochimie* **79**(6): 359-364.
- Asai, T., Bates, D.B., and Kogoma, T. 1994. DNA replication triggered by double-strand breaks in *E. coli*: dependence on homologous recombination functions. *Cell* **78**: 1051-1061.
- Au, K.G., Welsh, K., and Modrich, P. 1992. Initiation of methyl-directed mismatch repair. *J Biol Chem* **267**(17): 12142-12148.

- Aylon, Y. and Kupiec, M. 2004. DSB repair: the yeast paradigm. *DNA Repair (Amst)* **3**(8-9): 797-815.
- Azzaria, M., Schurr, E., and Gros, P. 1989. Discrete mutations introduced in the predicted nucleotide-binding sites of the *mdr1* gene abolish its ability to confer multidrug resistance. *Mol Cell Biol* **9**(12): 5289-5297.
- Bachmann, B.J. 1996. Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. in *Escherichia coli and Salmonella Cellular and Molecular Biology, (Second Edition)* (ed. F.C. Neidhardt, R. Curtiss III, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger), pp. 2460-2488. ASM Press, Washington, D.C.
- Barnes, D.E., Tomkinson, A.E., Lehmann, A.R., Webster, A.D.B., and Lindahl, T. 1992. Mutations in the DNA ligase I gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA-damaging agents. *Cell* **69**: 495-503.
- Bartosik, A.A. and Jagura-Burdzy, G. 2005. Bacterial chromosome segregation. *Acta Biochim Pol* **52**(1): 1-34.
- Beckman, K.B. and Ames, B.N. 1998. The free radical theory of aging matures. *Physiol Rev* **78**(2): 547-581.
- Belfort, M. and Roberts, R.J. 1997. Homing endonucleases: keeping the house in order. *Nucleic Acids Res* **25**(17): 3379-3388.
- Bender, A. and Pringle, J.R. 1991. Use of a screen for synthetic lethal and multicopy suppressor mutants to identify two new genes involved in morphogenesis in *Saccharomyces cerevisiae*. *Mol Cell Biol* **11**(3): 1295-1305.
- Bennett, R.J., Dunderdale, H.J., and West, S.C. 1993. Resolution of Holliday junctions by RuvC resolvase: cleavage specificity and DNA distortion. *Cell* **74**: 1021-1031.
- Berg, O.G. 1988. Selection of DNA binding sites by regulatory proteins : the LexA protein and the arginine repressor use different strategies for functional specificity. *Nucl Acids Res* **16**: 5089-5105.
- Bernhardt, T.G. and de Boer, P.A. 2003. The *Escherichia coli* amidase AmiC is a periplasmic septal ring component exported via the twin-arginine transport pathway. *Mol Microbiol* **48**(5): 1171-1182.
- . 2004. Screening for synthetic lethal mutants in *Escherichia coli* and identification of EnvC (YibP) as a periplasmic septal ring factor with murein hydrolase activity. *Mol Microbiol* **52**(5): 1255-1269.
- Bianco, P.R. and Kowalczykowski, S.C. 1997. The recombination hotspot Chi is recognized by the translocating RecBCD enzyme as the single strand of DNA containing the sequence 5'-GCTGGTGG- 3'. *Proc Natl Acad Sci U S A* **94**(13): 6706-6711.
- Bird, L.E., Hakansson, K., Pan, H., and Wigley, D.B. 1997. Characterization and crystallization of the helicase domain of bacteriophage T7 gene 4 protein. *Nucleic Acids Res* **25**(13): 2620-2626.
- Birkenbihl, R.P. and Subramani, S. 1992. Cloning and characterization of *rad21* an essential gene of *Schizosaccharomyces pombe* involved in DNA double-strand-break repair. *Nucleic Acids Res* **20**(24): 6605-6611.

- Bishop, B., Dasgupta, J., and Chen, X.S. 2007. Structure-based engineering of papillomavirus major capsid I1: controlling particle assembly. *Virol J* **4**: 3.
- Blank, M., Lerenthal, Y., Mittelman, L., and Shiloh, Y. 2006. Condensin I recruitment and uneven chromatin condensation precede mitotic cell death in response to DNA damage. *J Cell Biol* **174**(2): 195-206.
- Blat, Y. and Kleckner, N. 1999. Cohesins bind to preferential sites along yeast chromosome III, with differential regulation along arms versus the centric region. *Cell* **98**(2): 249-259.
- Boehmer, P.E. and Emmerson, P.T. 1991. *Escherichia coli* RecBCD enzyme: Inducible overproduction and reconstitution of the ATP-dependent deoxyribonuclease from purified subunits. *Gene* **102**: 1-6.
- Bonura, T. and Smith, K.C. 1975. Enzymatic production of deoxyribonucleic acid double-strand breaks after ultraviolet irradiation of *Escherichia coli* K-12. *J Bacteriol* **121**(2): 511-517.
- Bosco, G. and Haber, J.E. 1998. Chromosome break-induced DNA replication leads to nonreciprocal translocations and telomere capture. *Genetics* **150**(3): 1037-1047.
- Bradford, M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248.
- Bredeche, M.F., Ehrlich, S.D., and Michel, B. 2001. Viability of rep recA mutants depends on their capacity to cope with spontaneous oxidative damage and on the DnaK chaperone protein. *J Bacteriol* **183**(7): 2165-2171.
- Buonomo, S.B., Clyne, R.K., Fuchs, J., Loidl, J., Uhlmann, F., and Nasmyth, K. 2000. Disjunction of homologous chromosomes in meiosis I depends on proteolytic cleavage of the meiotic cohesin Rec8 by separin. *Cell* **103**(3): 387-398.
- Capaldo-Kimball, F. and Barbour, S.D. 1971. Involvement of recombination genes in growth and viability of *Escherichia coli* K-12. *J Bacteriol* **106**: 204-212.
- Chan, S.N., Harris, L., Bolt, E.L., Whitby, M.C., and Lloyd, R.G. 1997. Sequence-specificity and biochemical characterization of the RsaA Holliday junction resolvase of *Escherichia coli*. *J Biol Chem* **272**: 14873-14882.
- Chen, E.S., Sutani, T., and Yanagida, M. 2004. Ctl1/C1D interacts with condensin SMC hinge and supports the DNA repair function of condensin. *Proc Natl Acad Sci U S A* **101**(21): 8078-8083.
- Chen, N., Zinchenko, A.A., Yoshikawa, Y., Araki, S., Adachi, S., Yamazoe, M., Hiraga, S., and Yoshikawa, K. 2008. ATP-induced shrinkage of DNA with MukB protein and the MukBEF complex of *Escherichia coli*. *J Bacteriol* **190**(10): 3731-3737.
- Churchill, J.J. and Kowalczykowski, S.C. 2000. Identification of the RecA Protein-loading Domain of RecBCD Enzyme. *J Mol Biol* **297**(3): 537-542.
- Ciosk, R., Shirayama, M., Shevchenko, A., Tanaka, T., Toth, A., Shevchenko, A., and Nasmyth, K. 2000. Cohesin's binding to chromosomes depends

- on a separate complex consisting of Scc2 and Scc4 proteins. *Mol Cell* **5**(2): 243-254.
- Clark, A.J. and Margulies, A.D. 1965. Isolation and characterization of recombination deficient mutants of *Escherichia coli* K12. *Proc Natl Acad Sci USA* **53**: 451-459.
- Cobbe, N. and Heck, M.M. 2004. The evolution of SMC proteins: phylogenetic analysis and structural implications. *Molecular biology and evolution* **21**(2): 332-347.
- Cole, R.S. 1973. Repair of DNA containing interstrand crosslinks in *Escherichia coli*: Sequential excision and recombination. *Proc Natl Acad Sci USA* **70**: 1064-1068.
- Connelly, J.C., de Leau, E.S., and Leach, D.R. 1999. DNA cleavage and degradation by the SbcCD protein complex from *Escherichia coli*. *Nucleic Acids Res* **27**(4): 1039-1046.
- Cortes-Ledesma, F., de Piccoli, G., Haber, J.E., Aragon, L., and Aguilera, A. 2007. SMC proteins, new players in the maintenance of genomic stability. *Cell Cycle* **6**(8): 914-918.
- Cost, G.J. and Cozzarelli, N.R. 2006. Smc5p promotes faithful chromosome transmission and DNA repair in *Saccharomyces cerevisiae*. *Genetics* **172**(4): 2185-2200.
- Costantini, S., Woodbine, L., Andreoli, L., Jeggo, P.A., and Vindigni, A. 2007. Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK. *DNA Repair (Amst)* **6**(6): 712-722.
- Courcelle, J., Khodursky, A., Peter, B., Brown, P.O., and Hanawalt, P.C. 2001. Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. *Genetics* **158**(1): 41-64.
- Cox, M.M. 2001. Recombinational DNA repair of damaged replication forks in *Escherichia coli*: questions. *Annu Rev Genet* **35**: 53-82.
- Cox, M.M., Goodman, M.F., Kreuzer, K.N., Sherratt, D.J., Sandler, S.J., and Marians, K.J. 2000. The importance of repairing stalled replication forks. *Nature* **404**(6773): 37-41.
- Critchlow, S.E. and Jackson, S.P. 1998. DNA end-joining: from yeast to man. *Trends Biochem Sci* **23**(10): 394-398.
- Cromie, G.A. and Leach, D.R. 2001. Recombinational repair of chromosomal DNA double-strand breaks generated by a restriction endonuclease. *Mol Microbiol* **41**(4): 873-883.
- Cromie, G.A., Millar, C.B., Schmidt, K.H., and Leach, D.R. 2000. Palindromes as substrates for multiple pathways of recombination in *Escherichia coli*. *Genetics* **154**(2): 513-522.
- Datsenko, K.A. and Wanner, B.L. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* **97**(12): 6640-6645.
- Davies, A.A. and West, S.C. 1998. Formation of RuvABC-Holliday junction complexes *in vitro*. *Curr Biol* **8**(12): 725-727.
- De Piccoli, G., Cortes-Ledesma, F., Ira, G., Torres-Rosell, J., Uhle, S., Farmer, S., Hwang, J.Y., Machin, F., Ceschia, A., McAleenan, A., Cordon-Preciado, V., Clemente-Blanco, A., Vilella-Mitjana, F., Ullal, P.,

- Jarmuz, A., Leitao, B., Bressan, D., Dotiwala, F., Papusha, A., Zhao, X., Myung, K., Haber, J.E., Aguilera, A., and Aragon, L. 2006. Smc5-Smc6 mediate DNA double-strand-break repair by promoting sister-chromatid recombination. *Nat Cell Biol* **8**(9): 1032-1034.
- Deckert, G., Warren, P.V., Gaasterland, T., Young, W.G., Lenox, A.L., Graham, D.E., Overbeek, R., Snead, M.A., Keller, M., Aujay, M., Huber, R., Feldman, R.A., Short, J.M., Olsen, G.J., and Swanson, R.V. 1998. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. *Nature* **392**(6674): 353-358.
- DeFazio, L.G., Stansel, R.M., Griffith, J.D., and Chu, G. 2002. Synapsis of DNA ends by DNA-dependent protein kinase. *Embo J* **21**(12): 3192-3200.
- Derewenda, Z.S. 2004. The use of recombinant methods and molecular engineering in protein crystallization. *Methods* **34**(3): 354-363.
- Dillingham, M.S., Spies, M., and Kowalczykowski, S.C. 2003. RecBCD enzyme is a bipolar DNA helicase. *Nature* **423**(6942): 893-897.
- Doherty, A.J., Jackson, S.P., and Weller, G.R. 2001. Identification of bacterial homologues of the Ku DNA repair proteins. *FEBS Lett* **500**(3): 186-188.
- Donaldson, J.R., Courcelle, C.T., and Courcelle, J. 2004. RuvAB and RecG are not essential for the recovery of DNA synthesis following UV-induced DNA damage in *Escherichia coli*. *Genetics* **166**(4): 1631-1640.
- Dunman, P.M., Ren, L., Rahman, M.S., Palejwala, V.A., Murphy, H.S., Volkert, M.R., and Humayun, M.Z. 2000. *Escherichia coli* cells defective for the *recN* gene display constitutive elevation of mutagenesis at 3,N(4)-ethenocytosine via an SOS-induced mechanism. *Mol Microbiol* **37**(3): 680-686.
- Dye, K. and Ahmad, S.I. 1995. Repair of phage lambda DNA damaged by near ultraviolet light plus 8-methoxypsoralen. *J Gen Virol* **76** ( Pt 3): 723-726.
- Emmerson, P.T. and Howard-Flanders, P. 1965. Post-Irradiation Degradation Of Dna Following Exposure Of Uv-Sensitive And Resistant Bacteria To X-Rays. *Biochem Biophys Res Commun* **18**: 24-29.
- Erill, I., Campoy, S., and Barbe, J. 2007. Aeons of distress: an evolutionary perspective on the bacterial SOS response. *FEMS microbiology reviews* **31**(6): 637-656.
- Erill, I., Escribano, M., Campoy, S., and Barbe, J. 2003. In silico analysis reveals substantial variability in the gene contents of the gamma proteobacteria LexA-regulon. *Bioinformatics* **19**(17): 2225-2236.
- Eykelenboom, J.K., Blackwood, J.K., Okely, E., and Leach, D.R. 2008. SbcCD causes a double-strand break at a DNA palindrome in the *Escherichia coli* chromosome. *Mol Cell* **29**(5): 644-651.
- Fennell-Fezzie, R., Gradia, S.D., Akey, D., and Berger, J.M. 2005. The MukF subunit of *Escherichia coli* condensin: architecture and functional relationship to kleisins. *EMBO J* **24**(11): 1921-1930.
- Fernandez De Henestrosa, A.R., Ogi, T., Aoyagi, S., Chafin, D., Hayes, J.J., Ohmori, H., and Woodgate, R. 2000. Identification of additional genes

- belonging to the LexA regulon in *Escherichia coli*. *Mol Microbiol* **35**(6): 1560-1572.
- Finch, P.W., Chambers, P., and Emmerson, P.T. 1985a. Identification of the *E. coli* *recN* gene product as a major SOS protein. *J Bacteriol* **164**: 653-658.
- Fishel, R. and Kolodner, R. 1995. Identification of mismatch repair genes and their roles in the development of cancer. *Curr Opin Gen Dev* **5**: 382-395.
- Flynn, J.M., Levchenko, I., Sauer, R.T., and Baker, T.A. 2004. Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation. *Genes Dev* **18**(18): 2292-2301.
- Fousteri, M.I. and Lehmann, A.R. 2000. A novel SMC protein complex in *Schizosaccharomyces pombe* contains the Rad18 DNA repair protein. *Embo J* **19**(7): 1691-1702.
- Frelet, A. and Klein, M. 2006. Insight in eukaryotic ABC transporter function by mutation analysis. *FEBS Lett* **580**(4): 1064-1084.
- Friedberg, E.C. 1985. *DNA repair*. W.H. Freeman and Company, New York.
- . 2008. A brief history of the DNA repair field. *Cell Res* **18**(1): 3-7.
- Friedberg, E.C., Walker, G.C., and Siede, W. 1995. *DNA repair and mutagenesis*. ASM Press, Washington, D.C.
- Funayama, T., Narumi, I., Kikuchi, M., Kitayama, S., Watanabe, H., and Yamamoto, K. 1999. Identification and disruption analysis of the *recN* gene in the extremely radioresistant bacterium *Deinococcus radiodurans*. *Mutat Res* **435**(2): 151-161.
- Galletto, R., Amitani, I., Baskin, R.J., and Kowalczykowski, S.C. 2006. Direct observation of individual RecA filaments assembling on single DNA molecules. *Nature* **443**(7113): 875-878.
- Giraud-Panis, M.J. and Lilley, D.M. 1998. Structural recognition and distortion by the DNA junction-resolving enzyme RusA. *J Mol Biol* **278**(1): 117-133.
- Giraud, A., Radman, M., Matic, I., and Taddei, F. 2001. The rise and fall of mutator bacteria. *Curr Opin Microbiol* **4**(5): 582-585.
- Glynn, E.F., Megee, P.C., Yu, H.G., Mistrot, C., Unal, E., Koshland, D.E., DeRisi, J.L., and Gerton, J.L. 2004. Genome-wide mapping of the cohesin complex in the yeast *Saccharomyces cerevisiae*. *PLoS Biol* **2**(9): E259.
- Gong, C., Bongiorno, P., Martins, A., Stephanou, N.C., Zhu, H., Shuman, S., and Glickman, M.S. 2005. Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. *Nat Struct Mol Biol* **12**(4): 304-312.
- Goodman, M.F. 2002. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu Rev Biochem* **71**: 17-50.
- Gorbalenya, A.E. and Koonin, E.V. 1990. Superfamily of *uvrA*-related NTP-binding proteins. Implications for rational classification of recombination/repair systems. *J Mol Biol* **213**: 583-591.
- Gorbulev, S., Abele, R., and Tampe, R. 2001. Allosteric crosstalk between peptide-binding, transport, and ATP hydrolysis of the ABC transporter TAP. *Proc Natl Acad Sci U S A* **98**(7): 3732-3737.

- Graumann, P.L. 2001. SMC proteins in bacteria: condensation motors for chromosome segregation? *Biochimie* **83**(1): 53-59.
- Gregg, A.V., McGlynn, P., Jaktaji, R.P., and Lloyd, R.G. 2002. Direct rescue of stalled DNA replication forks via the combined action of PriA and RecG helicase activities. *Mol Cell* **9**(2): 241-251.
- Grove, J.I., Harris, L., Buckman, C., and Lloyd, R.G. 2008. DNA double strand break repair and crossing over mediated by RuvABC resolvase and RecG translocase. *DNA Repair (Amst)* **7**: 1517-1530.
- Gruber, S., Arumugam, P., Katou, Y., Kuglitsch, D., Helmhart, W., Shirahige, K., and Nasmyth, K. 2006. Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. *Cell* **127**(3): 523-537.
- Gruber, S., Haering, C.H., and Nasmyth, K. 2003. Chromosomal cohesin forms a ring. *Cell* **112**(6): 765-777.
- Guacci, V. 2007. Sister chromatid cohesion: the cohesin cleavage model does not ring true. *Genes Cells* **12**(6): 693-708.
- Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177**(14): 4121-4130.
- Haber, J.E. 1995. *in vivo* biochemistry: physical monitoring of recombination induced by site-specific endonucleases. *BioEssays* **17**(7): 609-620.
- . 1998a. The many interfaces of Mre11. *Cell* **95**(5): 583-586.
- . 1998b. Mating-type gene switching in *Saccharomyces cerevisiae*. *Annu Rev Genet* **32**: 561-599.
- . 2000. Lucky breaks: analysis of recombination in *Saccharomyces*. *Mutat Res* **451**(1-2): 53-69.
- Haering, C.H., Lowe, J., Hochwagen, A., and Nasmyth, K. 2002. Molecular architecture of SMC proteins and the yeast cohesin complex. *Mol Cell* **9**(4): 773-788.
- Haering, C.H. and Nasmyth, K. 2003. Building and breaking bridges between sister chromatids. *Bioessays* **25**(12): 1178-1191.
- Haering, C.H., Schoffnegger, D., Nishino, T., Helmhart, W., Nasmyth, K., and Lowe, J. 2004. Structure and stability of cohesin's Smc1-kleisin interaction. *Mol Cell* **15**(6): 951-964.
- Handa, N., Ohashi, S., and Kobayashi, I. 1997. Clustering of chi sequence in *Escherichia coli* genome. *Microb Comp Genomics* **2**(4): 287-298.
- Harvey, S.H., Sheedy, D.M., Cuddihy, A.R., and O'Connell, M.J. 2004. Coordination of DNA damage responses via the Smc5/Smc6 complex. *Mol Cell Biol* **24**(2): 662-674.
- Heale, J.T., Ball, A.R., Jr., Schmiesing, J.A., Kim, J.S., Kong, X., Zhou, S., Hudson, D.F., Earnshaw, W.C., and Yokomori, K. 2006. Condensin I interacts with the PARP-1-XRCC1 complex and functions in DNA single-strand break repair. *Mol Cell* **21**(6): 837-848.
- Heitman, J., Zinder, N.D., and Model, P. 1989. Repair of the *Escherichia coli* chromosome after *in vivo* scission by the EcoRI endonuclease. *Proc Natl Acad Sci U S A* **86**(7): 2281-2285.
- Higashitani, N., Higashitani, A., and Horiuchi, K. 1995. SOS induction in *Escherichia coli* by single-stranded DNA of mutant filamentous phage:



- monitoring by cleavage of LexA repressor. *J Bacteriol* **177**(12): 3610-3612.
- Higgins, C.F. 1992. ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* **8**: 67-113.
- Hiraga, S. 2000. Dynamic localization of bacterial and plasmid chromosomes. *Annu Rev Genet* **34**: 21-59.
- Hirano, M., Anderson, D.E., Erickson, H.P., and Hirano, T. 2001. Bimodal activation of SMC ATPase by intra- and inter-molecular interactions. *Embo J* **20**(12): 3238-3250.
- Hirano, M. and Hirano, T. 1998a. ATP-dependent aggregation of single-stranded DNA by a bacterial SMC homodimer. *Embo J* **17**(23): 7139-7148.
- . 2002. Hinge-mediated dimerization of SMC protein is essential for its dynamic interaction with DNA. *Embo J* **21**(21): 5733-5744.
- . 2004. Positive and negative regulation of SMC-DNA interactions by ATP and accessory proteins. *Embo J* **23**(13): 2664-2673.
- . 2006. Opening closed arms: long-distance activation of SMC ATPase by hinge-DNA interactions. *Mol Cell* **21**(2): 175-186.
- Hirano, T. 1999. SMC-mediated chromosome mechanics: a conserved scheme from bacteria to vertebrates? *Genes Dev* **13**(1): 11-19.
- . 2002. The ABCs of SMC proteins: two-armed ATPases for chromosome condensation, cohesion, and repair. *Genes Dev* **16**(4): 399-414.
- . 2005a. Condensins: organizing and segregating the genome. *Curr Biol* **15**(7): R265-275.
- . 2005b. SMC proteins and chromosome mechanics: from bacteria to humans. *Philos Trans R Soc Lond B Biol Sci* **360**(1455): 507-514.
- . 2006. At the heart of the chromosome: SMC proteins in action. *Nat Rev Mol Cell Biol* **7**(5): 311-322.
- Hoeijmakers, J.H. 2001. Genome maintenance mechanisms for preventing cancer. *Nature* **411**(6835): 366-374.
- Holland, I.B. and Blight, M.A. 1999. ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. *J Mol Biol* **293**(2): 381-399.
- Holliday, R. 1964. A mechanism for gene conversion in fungi. *Genet Res Camb* **5**: 282-304.
- Hopfner, K.P., Karcher, A., Shin, D.S., Craig, L., Arthur, L.M., Carney, J.P., and Tainer, J.A. 2000. Structural biology of Rad50 ATPase: ATP-driven conformational control in DNA double-strand break repair and the ABC-ATPase superfamily. *Cell* **101**(7): 789-800.
- Hopfner, K.P. and Tainer, J.A. 2003. Rad50/SMC proteins and ABC transporters: unifying concepts from high-resolution structures. *Curr Opin Struct Biol* **13**(2): 249-255.
- Hrycyna, C.A., Ramachandra, M., Germann, U.A., Cheng, P.W., Pastan, I., and Gottesman, M.M. 1999. Both ATP sites of human P-glycoprotein are essential but not symmetric. *Biochemistry* **38**(42): 13887-13899.
- Huang, C.E., Milutinovich, M., and Koshland, D. 2005. Rings, bracelet or snaps: fashionable alternatives for SMC complexes. *Philos Trans R Soc Lond B Biol Sci* **360**(1455): 537-542.

- Hudson, D.F., Ohta, S., Freisinger, T., Macisaac, F., Sennels, L., Alves, F., Lai, F., Kerr, A., Rappsilber, J., and Earnshaw, W.C. 2008. Molecular and Genetic Analysis of Condensin Function in Vertebrate Cells. *Mol Biol Cell*.
- Hyde, S.C., Emsley, P., Hartshorn, M.J., Mimmack, M.M., Gileadi, U., Pearce, S.R., Gallagher, M.P., Gill, D.R., Hubbard, R.E., and Higgins, C.F. 1990. Structural model of ATP-binding proteins associated with cystic fibrosis, multidrug resistance and bacterial transport. *Nature* **346**(6282): 362-365.
- Hyrien, O. 2000. Mechanisms and consequences of replication fork arrest. *Biochimie* **82**(1): 5-17.
- Ivanov, D. and Nasmyth, K. 2005. A topological interaction between cohesin rings and a circular minichromosome. *Cell* **122**(6): 849-860.
- . 2007. A physical assay for sister chromatid cohesion in vitro. *Mol Cell* **27**(2): 300-310.
- Iyer, V.N. and Szybalski, W. 1963. A molecular mechanism of mitomycin action: linking of complementary DNA strands. *Proc Natl Acad Sci USA* **50**: 355-362.
- Jaenicke, E., Walsh, P.J., and Decker, H. 2003. Isolation and characterization of haemoporphin, an abundant haemolymph protein from *Aplysia californica*. *Biochem J* **375**(3): 681-688.
- Janion, C. 2001. Some aspects of the SOS response system--a critical survey. *Acta Biochim Pol* **48**(3): 599-610.
- Jeggo, P.A., Carr, A.M., and Lehmann, A.R. 1998. Splitting the ATM: distinct repair and checkpoint defects in ataxia-telangiectasia. *Trends Genet* **14**(8): 312-316.
- Johnson, R.D. and Jasin, M. 2001. Double-strand-break-induced homologous recombination in mammalian cells. *Biochem Soc Trans* **29**(Pt 2): 196-201.
- Jones, D.T. 1999. GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. *J Mol Biol* **287**(4): 797-815.
- Keeney, S., Giroux, C.N., and Kleckner, N. 1997. Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* **88**(3): 375-384.
- Kelner, A. 1949. Effect of Visible Light on the Recovery of *Streptomyces Griseus* Conidia from Ultra-violet Irradiation Injury. *Proc Natl Acad Sci U S A* **35**(2): 73-79.
- Khil, P.P. and Camerini-Otero, R.D. 2002. Over 1000 genes are involved in the DNA damage response of *Escherichia coli*. *Mol Microbiol* **44**(1): 89-105.
- Kidane, D. and Graumann, P.L. 2005. Intracellular protein and DNA dynamics in competent *Bacillus subtilis* cells. *Cell* **122**(1): 73-84.
- Kidane, D., Sanchez, H., Alonso, J.C., and Graumann, P.L. 2004. Visualization of DNA double-strand break repair in live bacteria reveals dynamic recruitment of *Bacillus subtilis* RecF, RecO and RecN proteins to distinct sites on the nucleoids. *Mol Microbiol* **52**(6): 1627-1639.

- Kim, D.Y. and Kim, K.K. 2005. Structure and function of HtrA family proteins, the key players in protein quality control. *Journal of biochemistry and molecular biology* **38**(3): 266-274.
- Kim, J.S., Krasieva, T.B., LaMorte, V., Taylor, A.M., and Yokomori, K. 2002a. Specific recruitment of human cohesin to laser-induced DNA damage. *J Biol Chem* **277**(47): 45149-45153.
- Kim, S.T., Xu, B., and Kastan, M.B. 2002b. Involvement of the cohesin protein, Smc1, in Atm-dependent and independent responses to DNA damage. *Genes Dev* **16**(5): 560-570.
- Kleckner, N., Bender, J., and Gottesman, S. 1991. Uses of transposons with emphasis on Tn10. *Methods Enzymol* **204**: 139-180.
- Klein, F., Mahr, P., Galova, M., Buonomo, S.B., Michaelis, C., Nairz, K., and Nasmyth, K. 1999. A central role for cohesins in sister chromatid cohesion, formation of axial elements, and recombination during yeast meiosis. *Cell* **98**(1): 91-103.
- Kogoma, T. 1996. Recombination by replication. *Cell* **85**: 625-627.
- . 1997. Stable DNA replication: Interplay between DNA replication, homologous recombination, and transcription. *Microbiol Molec Biol Rev* **61**: 212-238.
- Kohanski, M.A., Dwyer, D.J., Hayete, B., Lawrence, C.A., and Collins, J.J. 2007. A common mechanism of cellular death induced by bactericidal antibiotics. *Cell* **130**(5): 797-810.
- Kolodner, R. 1996. Biochemistry and genetics of eukaryotic mismatch repair. *Genes Dev* **10**: 1433-1442.
- Kolodner, R., Hall, S.D., and Luisi-DeLuca, C. 1994. Homologous pairing proteins encoded by the *Escherichia coli* *recE* and *recT* genes. *Mol Microbiol* **11**: 23-30.
- Konola, J.T., Sargent, K.E., and Gow, J.B. 2000. Efficient repair of hydrogen peroxide-induced DNA damage by *Escherichia coli* requires SOS induction of RecA and RuvA proteins. *Mutat Res* **459**(3): 187-194.
- Kooistra, J., Haijema, B.J., and Venema, G. 1993. The *Bacillus subtilis* addAB genes are fully functional in *Escherichia coli*. *Mol Microbiol* **7**(6): 915-923.
- Koop, A.H., Hartley, M.E., and Bourgeois, S. 1987. A low-copy-number vector utilizing beta-galactosidase for the analysis of gene control elements. *Gene* **52**(2-3): 245-256.
- Kornberg, A. and Baker, T.A. 1992. *DNA replication*. Freeman, W.H., New York.
- Kosa, J.L., Zdraveski, Z.Z., Currier, S., Marinus, M.G., and Essigmann, J.M. 2004. RecN and RecG are required for *Escherichia coli* survival of Bleomycin-induced damage. *Mutat Res* **554**(1-2): 149-157.
- Kowalczykowski, S.C. 2000. Initiation of genetic recombination and recombination-dependent replication [In Process Citation]. *Trends Biochem Sci* **25**(4): 156-165.
- Krasin, F. and Hutchinson, F. 1981. Repair of DNA double-strand breaks in *Escherichia coli* cells requires synthesis of proteins that can be induced by UV light. *Proc Natl Acad Sci U S A* **78**(6): 3450-3453.

- Kraus, E., Leung, W.Y., and Haber, J.E. 2001. Break-induced replication: a review and an example in budding yeast. *Proc Natl Acad Sci U S A* **98**(15): 8255-8262.
- Krogh, B.O. and Symington, L.S. 2004. Recombination proteins in yeast. *Annu Rev Genet* **38**: 233-271.
- Krokan, H.E., Standal, R., and Slupphaug, G. 1997. DNA glycosylases in the base excision repair of DNA. *Biochem J* **325** ( Pt 1): 1-16.
- Kullik, I., Stevens, J., Toledano, M.B., and Storz, G. 1995. Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for DNA binding and multimerization. *J Bacteriol* **177**(5): 1285-1291.
- Kushner, S.R., Nagaishi, H., Templin, A., and Clark, A.J. 1971. Genetic recombination in *Escherichia coli*: the role of Exonuclease I. *Proc Natl Acad Sci USA* **68**: 824-827.
- Kuzminov, A. 1993. RuvA, RuvB and RuvC proteins: Cleaning up after recombinational repairs in *Escherichia coli*. *Bioessays* **15**: 355-358.
- . 1995a. Collapse and repair of replication forks in *Escherichia coli*. *Mol Microbiol* **16**(3): 373-384.
- . 1995b. Instability of inhibited replication forks in *E. coli*. *BioEssays* **17**(733-741).
- . 2001. Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc Natl Acad Sci U S A* **98**(15): 8241-8246.
- Lam, S.T., Stahl, M.M., McMilin, K.D., and Stahl, F.W. 1974. Rec-mediated recombinational hot spot activity in bacteriophage lambda. II. A mutation which causes hot spot activity. *Genetics* **77**: 425-433.
- Lammens, A., Schele, A., and Hopfner, K.P. 2004. Structural biochemistry of ATP-driven dimerization and DNA-stimulated activation of SMC ATPases. *Curr Biol* **14**(19): 1778-1782.
- Lapinski, P.E., Neubig, R.R., and Raghavan, M. 2001. Walker A lysine mutations of TAP1 and TAP2 interfere with peptide translocation but not peptide binding. *J Biol Chem* **276**(10): 7526-7533.
- Lau, I.F., Filipe, S.R., Soballe, B., Okstad, O.A., Barre, F.X., and Sherratt, D.J. 2003. Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol Microbiol* **49**(3): 731-743.
- Lee, J.W., Blanco, L., Zhou, T., Garcia-Diaz, M., Bebenek, K., Kunkel, T.A., Wang, Z., and Povirk, L.F. 2004. Implication of DNA polymerase lambda in alignment-based gap filling for nonhomologous DNA end joining in human nuclear extracts. *J Biol Chem* **279**(1): 805-811.
- Lehmann, A.R. 2005. The role of SMC proteins in the responses to DNA damage. *DNA Repair (Amst)* **4**(3): 309-314.
- . 2006. Translesion synthesis in mammalian cells. *Exp Cell Res* **312**(14): 2673-2676.
- Lehmann, A.R., Walicka, M., Griffiths, D.J., Murray, J.M., Watts, F.Z., McCready, S., and Carr, A.M. 1995. The rad18 gene of *Schizosaccharomyces pombe* defines a new subgroup of the SMC superfamily involved in DNA repair. *Mol Cell Biol* **15**(12): 7067-7080.
- Lengronne, A., Katou, Y., Mori, S., Yokobayashi, S., Kelly, G.P., Itoh, T., Watanabe, Y., Shirahige, K., and Uhlmann, F. 2004. Cohesin relocation

- from sites of chromosomal loading to places of convergent transcription. *Nature* **430**(6999): 573-578.
- Lengronne, A., McIntyre, J., Katou, Y., Kanoh, Y., Hopfner, K.P., Shirahige, K., and Uhlmann, F. 2006. Establishment of sister chromatid cohesion at the *S. cerevisiae* replication fork. *Mol Cell* **23**(6): 787-799.
- Lewis, L.K., Harlow, G.R., Gregg-Jolly, L.A., and Mount, D.W. 1994. Identification of high affinity binding sites for LexA which define new DNA damage-inducible genes in *Escherichia coli*. *J Mol Biol* **241**(4): 507-523.
- Lindahl, T., Sedgwick, B., Sekiguchi, M., and Nakabeppu, Y. 1988. Regulation and expression of the adaptive response to alkylating agents. *Ann Rev Biochem* **57**: 133-157.
- Lindahl, T. and Wood, R.D. 1999. Quality control by DNA repair. *Science* **286**(5446): 1897-1905.
- Lindroos, H.B., Strom, L., Itoh, T., Katou, Y., Shirahige, K., and Sjogren, C. 2006. Chromosomal association of the Smc5/6 complex reveals that it functions in differently regulated pathways. *Mol Cell* **22**(6): 755-767.
- Linn, S. and Imlay, J.A. 1987. Toxicity, mutagenesis and stress responses induced in *Escherichia coli* by hydrogen peroxide. *J Cell Sci Suppl* **6**: 289-301.
- Little, J.W., Mount, D.W., and Yanisch-Perron, C.R. 1981. Purified *lexA* protein is a repressor of the *recA* and *lexA* genes. *Proc Natl Acad Sci USA* **78**: 4199-4203.
- Liu, J. and Marians, K.J. 1999. PriA-directed assembly of a primosome on D loop DNA. *J Biol Chem* **274**(35): 25033-25041.
- Liu, J., Xu, L., Sandler, S.J., and Marians, K.J. 1999. Replication fork assembly at recombination intermediates is required for bacterial growth. *Proc Natl Acad Sci USA* **96**(7): 3552-3555.
- Lloyd, R.G. 1991. Conjugational recombination in resolvase-deficient *ruvC* mutants of *Escherichia coli* K-12 depends on *recG*. *J Bacteriol* **173**: 5414-5418.
- Lloyd, R.G. and Buckman, C. 1985. Identification and genetic analysis of *sbcC* mutations in commonly used *recBC sbcB* strains of *Escherichia coli* K-12. *J Bacteriol* **164**: 836-844.
- . 1991. Overlapping functions of *recD*, *recJ* and *recN* provide evidence of three epistatic groups of genes in *Escherichia coli* recombination and DNA repair. *Biochimie* **73**: 313-320.
- . 1995. Conjugational recombination in *Escherichia coli*: Genetic analysis of recombinant formation in Hfr x F<sup>-</sup> crosses. *Genetics* **139**: 1123-1148.
- Lloyd, R.G. and Low, K.B. 1996. Homologous recombination. in *Escherichia coli and Salmonella Cellular and Molecular Biology, (Second Edition)* (ed. F.C. Neidhardt, R. Curtiss III, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger), pp. 2236-2255. ASM Press, Washington, D.C.
- Lloyd, R.G., Picksley, S.M., and Prescott, C. 1983. Inducible expression of a gene specific to the RecF pathway for recombination in *Escherichia coli* K12. *Mol Gen Genet* **190**: 162-167.

- Lloyd, R.G., Porton, M.C., and Buckman, C. 1988. Effect of *recF*, *recJ*, *recN*, *recO* and *ruv* mutations on ultraviolet survival and genetic recombination in a *recD* strain of *Escherichia coli* K-12. *Mol Gen Genet* **212**: 317-324.
- Lloyd, R.G. and Sharples, G.J. 1993a. Dissociation of synthetic Holliday junctions by *E. coli* RecG protein. *EMBO J* **12**: 17-22.
- . 1993b. Dissociation of synthetic Holliday junctions by *E. coli* RecG protein. *EMBO J* **12**(1): 17-22.
- . 1993c. Processing of recombination intermediates by the RecG and RuvAB proteins of *Escherichia coli*. *Nucl Acids Res* **21**: 1719-1725.
- Lockhart, A. and Kendrick-Jones, J. 1998. Nucleotide-dependent interaction of the N-terminal domain of MukB with microtubules. *J Struct Biol* **124**(2-3): 303-310.
- Lombard, D.B., Chua, K.F., Mostoslavsky, R., Franco, S., Gostissa, M., and Alt, F.W. 2005. DNA repair, genome stability, and aging. *Cell* **120**(4): 497-512.
- Loo, T.W. and Clarke, D.M. 1995. Rapid purification of human P-glycoprotein mutants expressed transiently in HEK 293 cells by nickel-chelate chromatography and characterization of their drug-stimulated ATPase activities. *J Biol Chem* **270**(37): 21449-21452.
- Losada, A. and Hirano, T. 2005. Dynamic molecular linkers of the genome: the first decade of SMC proteins. *Genes Dev* **19**(11): 1269-1287.
- Lowe, J., Cordell, S.C., and van den Ent, F. 2001. Crystal structure of the SMC head domain: an ABC ATPase with 900 residues antiparallel coiled-coil inserted. *J Mol Biol* **306**(1): 25-35.
- Lukacsovich, T., Yang, D., and Waldman, A.S. 1994. Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-SceI. *Nucleic Acids Res* **22**(25): 5649-5657.
- Lupas, A., Van Dyke, M., and Stock, J. 1991. Predicting coiled coils from protein sequences. *Science* **252**: 1162-1164.
- Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M.R. 2002. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* **108**(6): 781-794.
- Mahajan, K.N., Nick McElhinny, S.A., Mitchell, B.S., and Ramsden, D.A. 2002. Association of DNA polymerase mu (pol mu) with Ku and ligase IV: role for pol mu in end-joining double-strand break repair. *Mol Cell Biol* **22**(14): 5194-5202.
- Mahdi, A.A., Sharples, G.J., Mandal, T.N., and Lloyd, R.G. 1996. Holliday junction resolvases encoded by homologous *rusA* genes in *Escherichia coli* K-12 and phage 82. *J Mol Biol* **257**: 561-573.
- Malkova, A., Ivanov, E.L., and Haber, J.E. 1996. Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication. *Proc Natl Acad Sci U S A* **93**(14): 7131-7136.
- Mandal, T.N., Mahdi, A.A., Sharples, G.J., and Lloyd, R.G. 1993. Resolution of Holliday intermediates in recombination and DNA repair: indirect suppression of *ruvA*, *ruvB* and *ruvC* mutations. *J Bacteriol* **175**: 4325-4334.

- Mari, P.O., Florea, B.I., Persengiev, S.P., Verkaik, N.S., Bruggenwirth, H.T., Modesti, M., Giglia-Mari, G., Bezstarosti, K., Demmers, J.A., Luiders, T.M., Houtsmuller, A.B., and van Gent, D.C. 2006. Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4. *Proc Natl Acad Sci U S A* **103**(49): 18597-18602.
- Marians, K.J. 2000. Replication and recombination intersect. *Curr Opin Genet Dev* **10**: 151-156.
- . 2004. Mechanisms of replication fork restart in Escherichia coli. *Philos Trans R Soc Lond B Biol Sci* **359**(1441): 71-77.
- Mascarenhas, J., Sanchez, H., Tadesse, S., Kidane, D., Krisnamurthy, M., Alonso, J.C., and Graumann, P.L. 2006. Bacillus subtilis SbcC protein plays an important role in DNA inter-strand cross-link repair. *BMC Mol Biol* **7**: 20.
- McGlynn, P., Al-Deib, A.A., Liu, J., Marians, K.J., and Lloyd, R.G. 1997. The DNA replication protein PriA and the recombination protein RecG bind D-loops. *J Mol Biol* **270**: 212-221.
- McGlynn, P. and Lloyd, R.G. 2000. Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression. *Cell* **101**: 35-45.
- . 2001. Rescue of stalled replication forks by RecG: simultaneous translocation on the leading and lagging strand templates supports an active DNA unwinding model of fork reversal and Holliday junction formation. *Proc Nat Acad Sci USA* **98**: 8227-8234.
- . 2002. Recombinational repair and restart of damaged replication forks. *Nat Rev Mol Cell Biol* **3**(11): 859-870.
- Mcglynn, P.a.L., R.G. 2000. Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression. *Cell* **101**: 35-45
- McGowan, C.H. and Russell, P. 2004. The DNA damage response: sensing and signaling. *Curr Opin Cell Biol* **16**(6): 629-633.
- McHugh, P.J., Spanswick, V.J., and Hartley, J.A. 2001. Repair of DNA interstrand crosslinks: molecular mechanisms and clinical relevance. *Lancet Oncol* **2**(8): 483-490.
- McIlwraith, M.J., Vaisman, A., Liu, Y., Fanning, E., Woodgate, R., and West, S.C. 2005. Human DNA polymerase eta promotes DNA synthesis from strand invasion intermediates of homologous recombination. *Mol Cell* **20**(5): 783-792.
- McIlwraith, M.J., Van Dyck, E., Masson, J.Y., Stasiak, A.Z., Stasiak, A., and West, S.C. 2000. Reconstitution of the strand invasion step of double-strand break repair using human rad51 rad52 and RPA proteins. *J Mol Biol* **304**: 151-164.
- McIlwraith, M.J. and West, S.C. 2008. DNA Repair Synthesis Facilitates RAD52-Mediated Second-End Capture during DSB Repair. *Mol Cell* **29**(4): 510-516.
- Meddows, T.R. 2002. DNA breakage and repair in Escherichia coli. In *Institute of genetics*, pp. 172. University of Nottingham.

- Meddows, T.R., Savory, A.P., Grove, J.I., Moore, T., and Lloyd, R.G. 2005. RecN protein and transcription factor DksA combine to promote faithful recombinational repair of DNA double-strand breaks. *Mol Microbiol* **57**(1): 97-110.
- Meddows, T.R., Savory, A.P., and Lloyd, R.G. 2004. RecG helicase promotes DNA double-strand break repair. *Mol Microbiol* **52**(1): 119-132.
- Melby, T.E., Ciampaglio, C.N., Briscoe, G., and Erickson, H.P. 1998. The symmetrical structure of structural maintenance of chromosomes (SMC) and MukB proteins: long, antiparallel coiled coils, folded at a flexible hinge. *J Cell Biol* **142**(6): 1595-1604.
- Meng, F.G., Hong, Y.K., He, H.W., Lyubarev, A.E., Kurganov, B.I., Yan, Y.B., and Zhou, H.M. 2004. Osmophobic effect of glycerol on irreversible thermal denaturation of rabbit creatine kinase. *Biophys J* **87**(4): 2247-2254.
- Mengiste, T., Revenkova, E., Bechtold, N., and Paszkowski, J. 1999. An SMC-like protein is required for efficient homologous recombination in Arabidopsis. *Embo J* **18**(16): 4505-4512.
- Meselson, M. and Weigle, J.J. 1961. Chromosome breakage accompanying genetic recombination in bacteriophage. *Proc Natl Acad Sci U S A* **47**: 857-868.
- Michel, B. 2000. Replication fork arrest and DNA recombination. *Trends Biochem Sci* **25**(4): 173-178.
- Michel, B., Ehrlich, S.D., and Uzest, M. 1997. DNA double-strand breaks caused by replication arrest. *EMBO J* **16**: 430-438.
- Michel, B., Flores, M.J., Viguera, E., Grompone, G., Seigneur, M., and Bidnenko, V. 2001. Rescue of arrested replication forks by homologous recombination. *Proc Natl Acad Sci U S A* **98**(15): 8181-8188.
- Miller, J.H. 1972. *Experiments in molecular genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Mirzoeva, O.K. and Petrini, J.H. 2001. DNA damage-dependent nuclear dynamics of the Mre11 complex. *Mol Cell Biol* **21**(1): 281-288.
- Mitchell, A.H. and West, S.C. 1994. Hexameric rings of *E. coli* RuvB protein: Cooperative assembly, processivity, and ATPase activity. *J Mol Biol* **243**: 208-215.
- Modrich, P. 1991. Mechanisms and biological effects of mismatch repair. *Annual Reviews in Genetics* **25**: 229-253.
- . 1994. Mismatch repair, genetic stability, and cancer. *Science* **266**(5193): 1959-1960.
- Moldovan, G.L., Pfander, B., and Jentsch, S. 2006. PCNA controls establishment of sister chromatid cohesion during S phase. *Mol Cell* **23**(5): 723-732.
- Moore, T. I, Unpublished data.
- Morimatsu, K. and Kowalczykowski, S.C. 2003. RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Mol Cell* **11**(5): 1337-1347.
- Morishita, T., Tsutsui, Y., Iwasaki, H., and Shinagawa, H. 2002. The *Schizosaccharomyces pombe* rad60 gene is essential for repairing double-strand DNA breaks spontaneously occurring during replication



- and induced by DNA-damaging agents. *Mol Cell Biol* **22**(10): 3537-3548.
- Moynahan, M.E., Chiu, J.W., Koller, B.H., and Jasin, M. 1999. BRCA1 controls homology-directed DNA repair. *Mol Cell* **4**: 511 - 518.
- Moynahan, M.E., Pierce, A.J., and Jasin, M. 2001. BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol Cell* **7**: 263 - 272.
- Mukherjee, A., Cao, C., and Lutkenhaus, J. 1998. Inhibition of FtsZ polymerization by SulA, an inhibitor of septation in Escherichia coli. *Proc Natl Acad Sci U S A* **95**(6): 2885-2890.
- Mukhopadhyay, S. and Schellhorn, H.E. 1997. Identification and characterization of hydrogen peroxide-sensitive mutants of Escherichia coli: genes that require OxyR for expression. *J Bacteriol* **179**(2): 330-338.
- Mullis, K.B. and Faloona, F.A. 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods in Enzymology* **155**: 335-350.
- Nagashima, K., Kubota, Y., Shibata, T., Sakaguchi, C., Shinagawa, H., and Hishida, T. 2006. Degradation of Escherichia coli RecN aggregates by ClpXP protease and its implications for DNA damage tolerance. *J Biol Chem* **281**(41): 30941-30946.
- Nasim, A. and Smith, B.P. 1975. Genetic control of radiation sensitivity in Schizosaccharomyces pombe. *Genetics* **79**(4): 573-582.
- Nasmyth, K. and Haering, C.H. 2005. The structure and function of SMC and kleisin complexes. *Annu Rev Biochem* **74**: 595-648.
- Neher, S.B., Flynn, J.M., Sauer, R.T., and Baker, T.A. 2003. Latent ClpX-recognition signals ensure LexA destruction after DNA damage. *Genes Dev* **17**(9): 1084-1089.
- Neher, S.B., Villen, J., Oakes, E.C., Bakalarski, C.E., Sauer, R.T., Gygi, S.P., and Baker, T.A. 2006. Proteomic profiling of ClpXP substrates after DNA damage reveals extensive instability within SOS regulon. *Mol Cell* **22**(2): 193-204.
- Neubauer, P., Hofmann, K., Holst, O., Mattiasson, B., and Kruschke, P. 1992. Maximizing the expression of a recombinant gene in Escherichia coli by manipulation of induction time using lactose as inducer. *Appl Microbiol Biotechnol* **36**(6): 739-744.
- Nick McElhinny, S.A., Snowden, C.M., McCarville, J., and Ramsden, D.A. 2000. Ku recruits the XRCC4-ligase IV complex to DNA ends. *Mol Cell Biol* **20**(9): 2996-3003.
- Niki, H., Imamura, R., Kitaoka, M., Yamanaka, K., Ogura, T., and Hiraga, S. 1992. E.coli MukB protein involved in chromosome partition forms a homodimer with a rod-and-hinge structure having DNA binding and ATP/GTP binding activities. *EMBO J* **11**(13): 5101-5109.
- Ochsner, U.A., Vasil, M.L., Alsabbagh, E., Parvatiyar, K., and Hassett, D.J. 2000. Role of the pseudomonas aeruginosa oxyR-recG operon in oxidative stress defense and DNA repair: OxyR-dependent regulation of katB-ankB, ahpB, and ahpC-ahpF [In Process Citation]. *J Bacteriol* **182**(16): 4533-4544.

- Offer, G., Hicks, M.R., and Woolfson, D.N. 2002. Generalized Crick equations for modeling noncanonical coiled coils. *J Struct Biol* **137**(1-2): 41-53.
- Ohsumi, K., Yamazoe, M., and Hiraga, S. 2001. Different localization of SeqA-bound nascent DNA clusters and MukF-MukE- MukB complex in *Escherichia coli* cells. *Mol Microbiol* **40**(4): 835-845.
- Palecek, J., Vidot, S., Feng, M., Doherty, A.J., and Lehmann, A.R. 2006. The Smc5-Smc6 DNA repair complex. bridging of the Smc5-Smc6 heads by the KLEISIN, Nse4, and non-Kleisin subunits. *J Biol Chem* **281**(48): 36952-36959.
- Paques, F. and Haber, J.E. 1999. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* **63**(2): 349-404.
- Parsons, C.A., Stasiak, A., Bennett, R.J., and West, S.C. 1995. Structure of a multisubunit complex that promotes DNA branch migration. *Nature* **374**: 375-378.
- Pebernard, S., McDonald, W.H., Pavlova, Y., Yates, J.R., 3rd, and Boddy, M.N. 2004. Nse1, Nse2, and a novel subunit of the Smc5-Smc6 complex, Nse3, play a crucial role in meiosis. *Mol Biol Cell* **15**(11): 4866-4876.
- Peterson, K.R. and Mount, D.W. 1993. Analysis of the genetic requirements for viability of *Escherichia coli* K-12 DNA adenine methylase (*dam*) mutants. *J Bacteriol* **175**: 7505-7508.
- Petrushenko, Z.M., Lai, C.H., Rai, R., and Rybenkov, V.V. 2006a. DNA reshaping by MukB. Right-handed knotting, left-handed supercoiling. *J Biol Chem* **281**(8): 4606-4615.
- Petrushenko, Z.M., Lai, C.H., and Rybenkov, V.V. 2006b. Antagonistic interactions of kleisins and DNA with bacterial Condensin MukB. *J Biol Chem* **281**(45): 34208-34217.
- Phipps, J., Nasim, A., and Miller, D.R. 1985. Recovery, repair, and mutagenesis in *Schizosaccharomyces pombe*. *Adv Genet* **23**: 1-72.
- Picksley, S.M., Attfield, P.V., and Lloyd, R.G. 1984a. Repair of DNA double-strand breaks in *Escherichia coli* K12 requires a functional *recN* product. *Mol Gen Genet* **195**: 267-274.
- Picksley, S.M., Lloyd, R.G., and Buckman, C. 1984b. Genetic analysis and regulation of inducible recombination in *Escherichia coli* K-12. *Cold Spring Harbor Symp Quant Biol* **49**: 469-474.
- Picksley, S.M., Morton, S.J., and Lloyd, R.G. 1985a. Molecular analysis of DNA repair and recombination in *Escherichia coli*: Identification of the *recN* product. In *Society for General Microbiology (104th meeting)*, Nottingham.
- . 1985b. The *recN* locus of *Escherichia coli* K12: molecular analysis and identification of the gene product. *Mol Gen Genet* **201**: 301-307.
- Pierce, A.J., Stark, J.M., Araujo, F.D., Moynahan, M.E., Berwick, M., and Jasin, M. 2001. Double-strand breaks and tumorigenesis. *Trends Cell Biol* **11**(11): S52-59.
- Pollard, E.C. and Achey, P.M. 1975. Induction of radioresistance in *Escherichia coli*. *Biophys J* **15**(11): 1141-1154.

- Poole, L.B. 2005. Bacterial defenses against oxidants: mechanistic features of cysteine-based peroxidases and their flavoprotein reductases. *Archives of biochemistry and biophysics* **433**(1): 240-254.
- Poole, L.B. and Ellis, H.R. 1996. Flavin-dependent alkyl hydroperoxide reductase from *Salmonella typhimurium*. 1. Purification and enzymatic activities of overexpressed AhpF and AhpC proteins. *Biochemistry* **35**(1): 56-64.
- Postow, L., Ullsperger, C., Keller, R.W., Bustamante, C., Vologodskii, A.V., and Cozzarelli, N.R. 2001. Positive torsional strain causes the formation of a four-way junction at replication forks. *J Biol Chem* **276**(4): 2790-2796.
- Potts, P.R., Porteus, M.H., and Yu, H. 2006. Human SMC5/6 complex promotes sister chromatid homologous recombination by recruiting the SMC1/3 cohesin complex to double-strand breaks. *Embo J* **25**(14): 3377-3388.
- Potts, P.R. and Yu, H. 2005. Human MMS21/NSE2 is a SUMO ligase required for DNA repair. *Mol Cell Biol* **25**(16): 7021-7032.
- . 2007. The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOylation of telomere-binding proteins. *Nat Struct Mol Biol* **14**(7): 581-590.
- Povirk, L.F., Wubter, W., Kohnlein, W., and Hutchinson, F. 1977. DNA double-strand breaks and alkali-labile bonds produced by bleomycin. *Nucleic Acids Res* **4**(10): 3573-3580.
- Quillardet, P., Rouffaud, M.A., and Bouige, P. 2003. DNA array analysis of gene expression in response to UV irradiation in *Escherichia coli*. *Res Microbiol* **154**(8): 559-572.
- Radman, M. 1974. *Phenomenology of an inducible mutagenic DNA repair pathway in Escherichia coli : SOS repair hypothesis*. Charles C. Thomas, Springfield, Illinois.
- . 1975. *SOS repair hypothesis : phenomenology of an inducible DNA repair which is accompanied by mutagenesis*. Plenum Publishing Corporation, New York.
- Rafferty, J.B., Sedelnikova, S.E., Hargreaves, D., Artymiuk, P.J., Baker, P.J., Sharples, G.J., Mahdi, A.A., Lloyd, R.G., and Rice, D.W. 1996. Crystal structure of DNA recombination protein RuvA and a model for its binding to the Holliday junction. *Science* **274**(5286): 415-421.
- Raleigh, E.A., Trimarchi, R., and Revel, H. 1989. Genetic and physical mapping of the mcrA (rglA) and mcrB (rglB) loci of *Escherichia coli* K-12. *Genetics* **122**(2): 279-296.
- Rao, B.J. and Radding, C.M. 1993. Homologous recognition promoted by RecA protein via non-Watson-Crick bonds between identical DNA strands. *Proc Natl Acad Sci USA* **90**: 6646-6650.
- Rinken, R. and Wackernagel, W. 1992. Inhibition of the recBCD-dependent activation of Chi recombination hot spots in SOS-induced cells of *Escherichia coli*. *J Bacteriol* **174**: 1172-1178.
- Roman, L.J., Eggleston, A.K., and Kowalczykowski, S.C. 1992. Processivity of the DNA helicase activity of *Escherichia coli* RecBCD enzyme. *J Biol Chem* **267**: 4207-4214.

- Rosen, B.P., Bhattacharjee, H., Zhou, T., and Walmsley, A.R. 1999. Mechanism of the ArsA ATPase. *Biochim Biophys Acta* **1461**(2): 207-215.
- Rostas, K., Morton, S.J., Picksley, S.M., and Lloyd, R.G. 1987. Nucleotide sequence and LexA regulation of the *Escherichia coli* *recN* gene. *Nucl Acids Res* **15**: 5041-5049.
- Rozen, F., Pelletier, J., Trachsel, H., and Sonenberg, N. 1989. A lysine substitution in the ATP-binding site of eucaryotic initiation factor 4A abrogates nucleotide-binding activity. *Mol Cell Biol* **9**(9): 4061-4063.
- Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. 1988. Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA Polymerase. *Science* **239**: 487-491.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. *Molecular cloning. A laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Samson, L. and Cairns, J. 1977. A new pathway of DNA repair in *Escherichia coli*. *Nature (New Biol)* **267**: 281-282.
- Sancar, A. 1996a. DNA excision repair. *Annu Rev Biochem* **65**: 43-81.
- . 1996b. No "End of History" for photolyases. *Science* **272**(5258): 48-49.
- Sancar, G.B. 2000. Enzymatic photoreactivation: 50 years and counting. *Mutat Res* **451**(1-2): 25-37.
- Sanchez, H. and Alonso, J.C. 2005. Bacillus subtilis RecN binds and protects 3'-single-stranded DNA extensions in the presence of ATP. *Nucleic Acids Res* **33**(7): 2343-2350.
- Sanchez, H., Cardenas, P.P., Yoshimura, S.H., Takeyasu, K., and Alonso, J.C. 2007a. Dynamic structures of Bacillus subtilis RecN DNA complexes. *Nucleic Acids Res*.
- Sanchez, H., Carrasco, B., Cozar, M.C., and Alonso, J.C. 2007b. Bacillus subtilis RecG branch migration translocase is required for DNA repair and chromosomal segregation. *Mol Microbiol* **65**(4): 920-935.
- Sanchez, H., Kidane, D., Castillo Cozar, M., Graumann, P.L., and Alonso, J.C. 2006. Recruitment of Bacillus subtilis RecN to DNA double-strand breaks in the absence of DNA end processing. *J Bacteriol* **188**(2): 353-360.
- Sandler, S.J. and Marians, K.J. 2000. Role of PriA in replication fork reactivation in *Escherichia coli*. *J Bacteriol* **182**: 9-13.
- Sargentini, N. and Smith, K.C. 1986. Quantitation of the involvement of the *recA*, *recB*, *recC*, *recF*, *recJ*, *recN*, *lexA*, *radA*, *radB*, *uvrD* and *umuC* genes in the repair of X-ray-induced DNA double-strand breaks in *Escherichia coli*. *Radiat Res* **107**: 58-72.
- Sargentini, N.J. and Smith, K.C. 1983. Characterization of an *Escherichia coli* mutant (*radB101*) sensitive to gamma and uv radiation, and methyl methanesulfonate. *Radiat Res* **93**: 461-478.
- . 1988. Genetic and phenotypic analyses indicating occurrence of the *recN262* and *radB101* mutations at the same locus in *Escherichia coli*. *J Bacteriol* **170**: 2392-2394.

- Savory, A.P. 2007. The Repair of Site-Specific Chromosomal Double-Strand Breaks in *Escherichia coli*. In *Institute of Genetics*, pp. 202. University of Nottingham.
- Schleiffer, A., Kaitna, S., Maurer-Stroh, S., Glotzer, M., Nasmyth, K., and Eisenhaber, F. 2003. Kleisins: a superfamily of bacterial and eukaryotic SMC protein partners. *Mol Cell* **11**(3): 571-575.
- Schmitt, L. and Tampe, R. 2002. Structure and mechanism of ABC transporters. *Curr Opin Struct Biol* **12**(6): 754-760.
- Schneider, E. and Hunke, S. 1998. ATP-binding-cassette (ABC) transport systems: functional and structural aspects of the ATP-hydrolyzing subunits/domains. *FEMS microbiology reviews* **22**(1): 1-20.
- Seaver, L.C. and Imlay, J.A. 2001a. Alkyl hydroperoxide reductase is the primary scavenger of endogenous hydrogen peroxide in *Escherichia coli*. *J Bacteriol* **183**(24): 7173-7181.
- . 2001b. Hydrogen peroxide fluxes and compartmentalization inside growing *Escherichia coli*. *J Bacteriol* **183**(24): 7182-7189.
- . 2004. Are respiratory enzymes the primary sources of intracellular hydrogen peroxide? *J Biol Chem* **279**(47): 48742-48750.
- Seigneur, M., Bidnenko, V., Ehrlich, S.D., and Michel, B. 1998. RuvAB acts at arrested replication forks. *Cell* **95**: 419-430.
- Seigneur, M., Ehrlich, S.D., and Michel, B. 2000. RuvABC-dependent double-strand breaks in *dnaBts* mutants require RecA. *Mol Microbiol* **38**(3): 565-574.
- Sergeant, J., Taylor, E., Palecek, J., Foustier, M., Andrews, E.A., Sweeney, S., Shinagawa, H., Watts, F.Z., and Lehmann, A.R. 2005. Composition and architecture of the *Schizosaccharomyces pombe* Rad18 (Smc5-6) complex. *Mol Cell Biol* **25**(1): 172-184.
- Shah, R., Cosstick, R., and West, S.C. 1997. The RuvC protein dimer resolves Holliday junctions by a dual incision mechanism that involves base-specific contacts. *EMBO J* **16**: 1464-1472.
- Sharples, G.J., Chan, S.C., Mahdi, A.A., Whitby, M.C., and Lloyd, R.G. 1994. Processing of intermediates in recombination and DNA repair: identification of a new endonuclease that specifically cleaves Holliday junctions. *EMBO J* **13**: 6133-6142.
- Sharples, G.J., Ingleston, S.M., and Lloyd, R.G. 1999. Holliday junction processing in bacteria: insights from the evolutionary conservation of RuvABC, RecG, and RusA. *J Bacteriol* **181**: 5543-5550.
- Sharples, G.J. and Leach, D.R.F. 1995. Structural and functional similarities between SbcCD proteins of *Escherichia coli* and the RAD50 and MRE11 (RAD32) proteins of yeast. *Molec Microbiol* **17**: 1215-1220.
- Sharples, G.J. and Lloyd, R.G. 1993. Location of the *Bacillus subtilis sbcD* gene downstream of *addAB*, the analogues of *E. coli recBC*. *Nucleic Acids Res* **21**: 2010.
- Shcherbakova, P.V. and Fijalkowska, I.J. 2006. Translesion synthesis DNA polymerases and control of genome stability. *Front Biosci* **11**: 2496-2517.

- Shinohara, A., Ogawa, H., and Ogawa, T. 1992. Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. *Cell* **69**: 457-470.
- Shintomi, K. and Hirano, T. 2007. How are cohesin rings opened and closed? *Trends Biochem Sci* **32**(4): 154-157.
- Shlomai, J. and Kornberg, A. 1980. An *Escherichia coli* replication protein that recognises a unique sequence with a hairpin region in  $\phi$ X174 DNA. *Proc Natl Acad Sci USA* **77**: 799-803.
- Shrivastav, M., De Haro, L.P., and Nickoloff, J.A. 2008. Regulation of DNA double-strand break repair pathway choice. *Cell Res* **18**(1): 134-147.
- Shuman, S. and Glickman, M.S. 2007. Bacterial DNA repair by non-homologous end joining. *Nat Rev Microbiol* **5**(11): 852-861.
- Siddiqi, M.A. and Bothe, E. 1987. Single- and double-strand break formation in DNA irradiated in aqueous solution: dependence on dose and OH radical scavenger concentration. *Radiat Res* **112**(3): 449-463.
- Simic, D., Vukovic-Gacic, B., Ajanovic, A., and Knezevic-Vukcevic, J. 1991. Activation of RecA protein in recombination-deficient strains of *Escherichia coli* following DNA-damaging treatments. *Mutat Res* **254**(3): 255-262.
- Singleton, M.R., Dillingham, M.S., Gaudier, M., Kowalczykowski, S.C., and Wigley, D.B. 2004. Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature* **432**(7014): 187-193.
- Sjogren, C. and Nasmyth, K. 2001. Sister chromatid cohesion is required for postreplicative double-strand break repair in *Saccharomyces cerevisiae*. *Curr Biol* **11**(12): 991-995.
- Skaar, E.P., Lazio, M.P., and Seifert, H.S. 2002. Roles of the *recJ* and *recN* genes in homologous recombination and DNA repair pathways of *Neisseria gonorrhoeae*. *J Bacteriol* **184**(4): 919-927.
- Skibbens, R.V., Corson, L.B., Koshland, D., and Hieter, P. 1999. Ctf7p is essential for sister chromatid cohesion and links mitotic chromosome structure to the DNA replication machinery. *Genes Dev* **13**(3): 307-319.
- Smith, P.C., Karpowich, N., Millen, L., Moody, J.E., Rosen, J., Thomas, P.J., and Hunt, J.F. 2002. ATP binding to the motor domain from an ABC transporter drives formation of a nucleotide sandwich dimer. *Mol Cell* **10**(1): 139-149.
- Sohal, R.S. 2002. Role of oxidative stress and protein oxidation in the aging process. *Free Radic Biol Med* **33**(1): 37-44.
- Sonoda, E., Matsusaka, T., Morrison, C., Vagnarelli, P., Hoshi, O., Ushiki, T., Nojima, K., Fukagawa, T., Waizenegger, I.C., Peters, J.M., Earnshaw, W.C., and Takeda, S. 2001. Scc1/Rad21/Mcd1 is required for sister chromatid cohesion and kinetochore function in vertebrate cells. *Dev Cell* **1**(6): 759-770.
- Spies, M., Dillingham, M.S., and Kowalczykowski, S.C. 2005. Translocation by the RecB motor is an absolute requirement for {chi}-recognition and RecA protein loading by RecBCD enzyme. *J Biol Chem* **280**(44): 37078-37087.

- Spies, M. and Kowalczykowski, S.C. 2006. The RecA binding locus of RecBCD is a general domain for recruitment of DNA strand exchange proteins. *Mol Cell* **21**(4): 573-580.
- Stasiak, A., Tsaneva, I.R., West, S.C., Benson, C.J.B., Yu, X., and Egelman, E.H. 1994. The *E. coli* RuvB branch migration protein forms double hexameric rings around DNA. *Proc Natl Acad Sci USA* **91**: 7618-7622.
- Stephanou, N.C., Gao, F., Bongiorno, P., Ehrt, S., Schnappinger, D., Shuman, S., and Glickman, M.S. 2007. Mycobacterial nonhomologous end joining mediates mutagenic repair of chromosomal double-strand DNA breaks. *J Bacteriol* **189**(14): 5237-5246.
- Stohl, E.A. and Seifert, H.S. 2006. Neisseria gonorrhoeae DNA recombination and repair enzymes protect against oxidative damage caused by hydrogen peroxide. *J Bacteriol* **188**(21): 7645-7651.
- Storz, G. and Imlay, J.A. 1999. Oxidative stress. *Curr Opin Microbiol* **2**(2): 188-194.
- Storz, G., Jacobson, F.S., Tartaglia, L.A., Morgan, R.W., Silveira, L.A., and Ames, B.N. 1989. An alkyl hydroperoxide reductase induced by oxidative stress in *Salmonella typhimurium* and *Escherichia coli* : genetic characterization and cloning of *ahp*. *J Bacteriol* **171**: 2049-2055.
- Strom, L., Karlsson, C., Lindroos, H.B., Wedahl, S., Katou, Y., Shirahige, K., and Sjogren, C. 2007. Postreplicative formation of cohesion is required for repair and induced by a single DNA break. *Science* **317**(5835): 242-245.
- Strom, L., Lindroos, H.B., Shirahige, K., and Sjogren, C. 2004. Postreplicative recruitment of cohesin to double-strand breaks is required for DNA repair. *Mol Cell* **16**(6): 1003-1015.
- Strom, L. and Sjogren, C. 2005. DNA damage-induced cohesion. *Cell Cycle* **4**(4): 536-539.
- . 2007. Chromosome segregation and double-strand break repair - a complex connection. *Curr Opin Cell Biol* **19**(3): 344-349.
- Studier, F.W. and Moffat, B. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189**: 113-130.
- Sugawara, N. and Haber, J.E. 2006. Repair of DNA double strand breaks: in vivo biochemistry. *Methods Enzymol* **408**: 416-429.
- Sugiyama, T., Kantake, N., Wu, Y., and Kowalczykowski, S.C. 2006. Rad52-mediated DNA annealing after Rad51-mediated DNA strand exchange promotes second ssDNA capture. *EMBO J* **25**(23): 5539-5548.
- Sugiyama, T. and Kowalczykowski, S.C. 2002. Rad52 protein associates with replication protein A (RPA)-single-stranded DNA to accelerate Rad51-mediated displacement of RPA and presynaptic complex formation. *J Biol Chem* **277**(35): 31663-31672.
- Sunako, Y., Onogi, T., and Hiraga, S. 2001. Sister chromosome cohesion of *Escherichia coli*. *Mol Microbiol* **42**(5): 1233-1241.
- Sweetman, W.A., Moxon, E.R., and Bayliss, C.D. 2005. Induction of the SOS regulon of *Haemophilus influenzae* does not affect phase variation rates at tetranucleotide or dinucleotide repeats. *Microbiology* **151**(Pt 8): 2751-2763.

- Szabo, K., Welker, E., Bakos, M., Roninson, I., Varadi, A., and Sarkadi, B. 1998. Drug-stimulated nucleotide trapping in the human multidrug transporter MDR1. Cooperation of the nucleotide binding domains. *J Biol Chem* **273**(17): 10132-10138.
- Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., and Stahl, F.W. 1983. The double-strand-break repair model for recombination. *Cell* **33**(1): 25-35.
- Tabor, S. and Richardson, C.C. 1985. A bacteriophage T7-RNA polymerase promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci USA* **82**: 1074-1078.
- Tanaka, T., Fuchs, J., Loidl, J., and Nasmyth, K. 2000. Cohesin ensures bipolar attachment of microtubules to sister centromeres and resists their precocious separation. *Nat Cell Biol* **2**(8): 492-499.
- Tartaglia, L.A., Storz, G., and Ames, B.N. 1989. Identification and molecular analysis of *oxyR*-regulated promoters important for the bacterial adaptation to oxidative stress. *J Mol Biol* **210**: 709-719.
- Tatebayashi, K., Kato, J., and Ikeda, H. 1998. Isolation of a *Schizosaccharomyces pombe* rad21ts mutant that is aberrant in chromosome segregation, microtubule function, DNA repair and sensitive to hydroxyurea: possible involvement of Rad21 in ubiquitin-mediated proteolysis. *Genetics* **148**(1): 49-57.
- Taylor, A.F., Schultz, D.W., Ponticelli, A.S., and Smith, G.R. 1985. RecBC enzyme nicking at Chi sites during unwinding: Location and orientation-dependence of the cutting. *Cell* **41**: 153-163.
- Taylor, A.F. and Smith, G.R. 1995. Strand specificity of nicking of DNA at Chi sites by RecBCD enzyme. Modulation by ATP and magnesium levels. *J Biol Chem* **270**(41): 24459-24467.
- . 2003. RecBCD enzyme is a DNA helicase with fast and slow motors of opposite polarity. *Nature* **423**(6942): 889-893.
- Tomonaga, T., Nagao, K., Kawasaki, Y., Furuya, K., Murakami, A., Morishita, J., Yuasa, T., Sutani, T., Kearsey, S.E., Uhlmann, F., Nasmyth, K., and Yanagida, M. 2000. Characterization of fission yeast cohesin: essential anaphase proteolysis of Rad21 phosphorylated in the S phase. *Genes Dev* **14**(21): 2757-2770.
- Tonkin, E.T., Wang, T.J., Lisgo, S., Bamshad, M.J., and Strachan, T. 2004. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. *Nat Genet* **36**(6): 636-641.
- Torres-Rosell, J., Machin, F., Farmer, S., Jarmuz, A., Eydmann, T., Dalgaard, J.Z., and Aragon, L. 2005. SMC5 and SMC6 genes are required for the segregation of repetitive chromosome regions. *Nat Cell Biol* **7**(4): 412-419.
- Toth, A., Ciosk, R., Uhlmann, F., Galova, M., Schleiffer, A., and Nasmyth, K. 1999. Yeast cohesin complex requires a conserved protein, Eco1p(Ctf7), to establish cohesion between sister chromatids during DNA replication. *Genes Dev* **13**(3): 320-333.
- Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T., and Sedgwick, B. 2002. Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature* **419**(6903): 174-178.



- Trivedi, S., Gehlot, H.S., and Rao, S.R. 2006. Protein thermostability in Archaea and Eubacteria. *Genet Mol Res* **5**(4): 816-827.
- Trusca, D., Scott, S., Thompson, C., and Bramhill, D. 1998. Bacterial SOS checkpoint protein Sula inhibits polymerization of purified FtsZ cell division protein. *J Bacteriol* **180**(15): 3946-3953.
- Tsaneva, I.R., Müller, B., and West, S.C. 1992. ATP-dependent branch migration of Holliday junctions promoted by the RuvA and RuvB proteins of *E. coli*. *Cell* **69**: 1171-1180.
- Tsuyama, T., Inou, K., Seki, M., Seki, T., Kumata, Y., Kobayashi, T., Kimura, K., Hanaoka, F., Enomoto, T., and Tada, S. 2006. Chromatin loading of Smc5/6 is induced by DNA replication but not by DNA double-strand breaks. *Biochem Biophys Res Commun* **351**(4): 935-939.
- Uematsu, N., Weterings, E., Yano, K., Morotomi-Yano, K., Jakob, B., Taucher-Scholz, G., Mari, P.O., van Gent, D.C., Chen, B.P., and Chen, D.J. 2007. Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks. *J Cell Biol* **177**(2): 219-229.
- Uhlmann, F., Wernic, D., Poupart, M.A., Koonin, E.V., and Nasmyth, K. 2000. Cleavage of cohesin by the CD clan protease separin triggers anaphase in yeast. *Cell* **103**(3): 375-386.
- Unal, E., Arbel-Eden, A., Sattler, U., Shroff, R., Lichten, M., Haber, J.E., and Koshland, D. 2004. DNA damage response pathway uses histone modification to assemble a double-strand break-specific cohesin domain. *Mol Cell* **16**(6): 991-1002.
- Unal, E., Heidinger-Pauli, J.M., and Koshland, D. 2007. DNA double-strand breaks trigger genome-wide sister-chromatid cohesion through Eco1 (Ctf7). *Science* **317**(5835): 245-248.
- Uribe, S. and Sampedro, J.G. 2003. Measuring Solution Viscosity and its Effect on Enzyme Activity. *Biol Proced Online* **5**: 108-115.
- Usui, T., Ogawa, H., and Petrini, J.H. 2001. A DNA damage response pathway controlled by Tel1 and the Mre11 complex. *Mol Cell* **7**(6): 1255-1266.
- Usui, T., Ohta, T., Oshiumi, H., Tomizawa, J., Ogawa, H., and Ogawa, T. 1998. Complex formation and functional versatility of Mre11 of budding yeast in recombination. *Cell* **95**(5): 705-716.
- Uzest, M., Ehrlich, S.D., and Michel, B. 1995. Lethality of *rep recB* and *rep recC* double mutants of *Escherichia coli*. *Mol Microbiol* **17**(6): 1177-1188.
- van den Ent, F., Lockhart, A., Kendrick-Jones, J., and Lowe, J. 1999. Crystal structure of the N-terminal domain of MukB: a protein involved in chromosome partitioning. *Structure* **7**(10): 1181-1187.
- van Gent, D.C., Hoeijmakers, J.H., and Kanaar, R. 2001. Chromosomal stability and the DNA double-stranded break connection. *Nat Rev Genet* **2**(3): 196-206.
- van Gool, A.J., Hajibagheri, N.M., Stasiak, A., and West, S.C. 1999. Assembly of the *Escherichia coli* RuvABC resolvase directs the orientation of Holliday junction resolution. *Genes Dev* **13**(14): 1861-1870.
- van Gool, A.J., Shah, R., Mezard, C., and West, S.C. 1998. Functional interactions between the holliday junction resolvase and the branch migration motor of *Escherichia coli*. *Embo J* **17**(6): 1838-1845.

- Van Hoy, B.E. and Hoch, J.A. 1990. Characterization of the spoIVB and recN loci of *Bacillus subtilis*. *J Bacteriol* **172**(3): 1306-1311.
- Verkade, H.M., Bugg, S.J., Lindsay, H.D., Carr, A.M., and O'Connell, M.J. 1999. Rad18 is required for DNA repair and checkpoint responses in fission yeast. *Mol Biol Cell* **10**(9): 2905-2918.
- Viitanen, P.V., Lubben, T.H., Reed, J., Goloubinoff, P., O'Keefe, D.P., and Lorimer, G.H. 1990. Chaperonin-facilitated refolding of ribulosebiphosphate carboxylase and ATP hydrolysis by chaperonin 60 (groEL) are K<sup>+</sup> dependent. *Biochemistry* **29**(24): 5665-5671.
- Volkert, M.R. and Landini, P. 2001. Transcriptional responses to DNA damage. *Curr Opin Microbiol* **4**(2): 178-185.
- Walker, G.C. 1996. The SOS response of *Escherichia coli*. in *Escherichia coli and Salmonella Cellular and Molecular Biology, (Second Edition)* (ed. F.C. Neidhardt, R. Curtiss III, J.L. Ingraham, E.C.C. Lin, K.B. Low, B. Magasanik, W.S. Reznikoff, M. Riley, M. Schaechter, and H.E. Umbarger), pp. 1400-1416. ASM Press, Washington, D.C.
- Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. 1982. Distantly related sequences in the  $\alpha$ - and  $\beta$ -subunits of ATP synthase, myosin, kinase and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J* **1**: 945-951.
- Walker, J.R., Corpina, R.A., and Goldberg, J. 2001. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412**(6847): 607-614.
- Wang, G. and Maier, R.J. 2008. Critical role of RecN in recombinational DNA repair and survival of *Helicobacter pylori*. *Infect Immun* **76**(1): 153-160.
- Wang, H.C., Zeng, Z.C., Bui, T.A., DiBiase, S.J., Qin, W., Xia, F., Powell, S.N., and Iliakis, G. 2001. Non-homologous end-joining of ionizing radiation-induced DNA double-stranded breaks in human tumor cells deficient in BRCA1 or BRCA2. *Cancer Res* **61**: 270 - 277.
- Wang, T.V. and Smith, K.C. 1988. Different effects of *recJ* and *recN* mutations on the postreplication repair of UV-damaged DNA in *Escherichia coli* K-12. *J Bacteriol* **170**: 2555-2559.
- Wang, X., Reyes-Lamothe, R., and Sherratt, D.J. 2008. Modulation of *Escherichia coli* sister chromosome cohesion by topoisomerase IV. *Genes & Development* **22**: 2246-2433.
- Wang, Z., Cummins, J.M., Shen, D., Cahill, D.P., Jallepalli, P.V., Wang, T.L., Parsons, D.W., Traverso, G., Awad, M., Silliman, N., Ptak, J., Szabo, S., Willson, J.K., Markowitz, S.D., Goldberg, M.L., Karess, R., Kinzler, K.W., Vogelstein, B., Velculescu, V.E., and Lengauer, C. 2004. Three classes of genes mutated in colorectal cancers with chromosomal instability. *Cancer Res* **64**(9): 2998-3001.
- Ward, J.F. 1975. Molecular mechanisms of radiation-induced damage to nucleic acids. *Adv Rad Biol* **5**: 181-239.
- . 1988. DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability. *Prog Nucleic Acid Res Mol Biol* **35**: 95-125.

- Weitzer, S., Lehane, C., and Uhlmann, F. 2003. A model for ATP hydrolysis-dependent binding of cohesin to DNA. *Curr Biol* **13**(22): 1930-1940.
- Weller, G.R., Kysela, B., Roy, R., Tonkin, L.M., Scanlan, E., Della, M., Devine, S.K., Day, J.P., Wilkinson, A., di Fagagna, F., Devine, K.M., Bowater, R.P., Jeggo, P.A., Jackson, S.P., and Doherty, A.J. 2002. Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* **297**(5587): 1686-1689.
- Wen, Q., Mahdi, A.A., Briggs, G.S., Sharples, G.J., and Lloyd, R.G. 2005. Conservation of RecG activity from pathogens to hyperthermophiles. *DNA Repair (Amst)* **4**(1): 23-31.
- Whitby, M.C., Bolt, E.L., Chan, S.N., and Lloyd, R.G. 1996. Interactions between RuvA and RuvC at Holliday junctions: inhibition of junction cleavage and formation of a RuvA-RuvC-DNA complex. *J Mol Biol* **264**: 878-890.
- Wickner, S. and Hurwitz, J. 1975. Association of  $\phi$ X174 DNA-dependent ATPase activity with an *Escherichia coli* protein, replication factor Y, required for *in vitro* synthesis of  $\phi$ X174 DNA. *Proc Natl Acad Sci USA* **72**: 3342-3346.
- Witkin, E.M. 1976. Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriol Rev* **40**: 869-907.
- Wood, R.D. 1996. DNA repair in eukaryotes. *Annu Rev Biochem* **65**: 135-167.
- Yaneva, M., Kowalewski, T., and Lieber, M.R. 1997. Interaction of DNA-dependent protein kinase with DNA and with Ku: biochemical and atomic-force microscopy studies. *Embo J* **16**(16): 5098-5112.
- Yu, M., Souaya, J., and Julin, D.A. 1998. The 30-kDa C-terminal domain of the RecB protein is critical for the nuclease activity, but not the helicase activity, of the RecBCD enzyme from *Escherichia coli*. *Proc Natl Acad Sci USA* **95**(3): 981-986.
- Yu, V., Koehler, M., Steinlein, C., Schmid, M., Hanakahi, L.A., van Gool, A.J., West, S.C., and Venkitaraman, A.R. 2000. Gross chromosomal rearrangements and genetic exchange between non-homologous chromosomes following BRCA2 inactivation. *Genes Dev* **14**: 1400 - 1406.
- Yu, X. and Egelman, E.H. 1992. Structural data suggest that the active and inactive forms of the recA filament are not simply interconvertible. *J Mol Biol* **227**: 334-346.
- Zhao, X. and Blobel, G. 2005. A SUMO ligase is part of a nuclear multiprotein complex that affects DNA repair and chromosomal organization. *Proc Natl Acad Sci USA* **102**(13): 4777-4782.
- Zheng, M., Aslund, F., and Storz, G. 1998. Activation of the OxyR transcription factor by reversible disulfide bond formation. *Science* **279**(5357): 1718-1721.